CODEN: JASMAN

The Journal of the Acoustical Society of America

ISSN: 0001-4966

Vol. 115, No. 5, Pt. 1 May 2004

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 115, NO. 5, PT. 1, MAY 2004

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 115, NO. 5, PT. 1, MAY 2004

CONTENTS—*Continued from preceding page*

(Continued)

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOL. 115, NO. 5, PT. 1, MAY 2004

CONTENTS—*Continued from preceding page*

ACOUSTICAL NEWS—USA

Elaine Moran
Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502

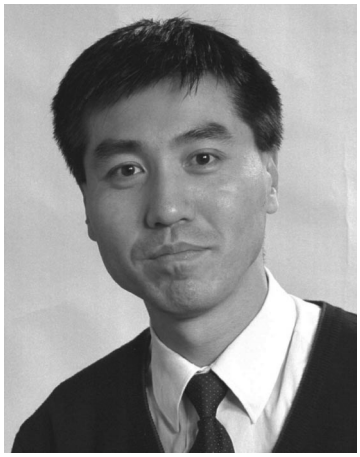
Editor's Note: *Readers of this Journal are encouraged to submit news items on awards, appointments, and other activities about themselves or their colleagues. Deadline dates for news items and notices are 2 months prior to publication.*

New Fellows of the Acoustical Society of America

David H. Chambers—For contributions to time reversal processing methodology.

John J. Guinan, Jr.—For contributions to the understanding of the physiology of auditory efferent feedback.

Gary J. Heald—For contributions to seabed classification by acoustical scattering methods.



Jian Kang—For contributions to teaching architectural acoustics and understanding complex urban noise environments.

Christopher J. Plack—For contributions to the understanding of temporal processing, intensity coding, and pitch perception

Ralph K. Hillquist named SAE Fellow

Ralph K. Hillquist of Benzonia, Michigan, has been elected to the status of ''SAE Fellow'' by the Society of Automotive Engineers (SAE International). Mr. Hillquist is a Technical Consultant and Vice President with RKH Consults and is a Fellow of the Acoustical Society of America.

Mr. Hillquist has made significant contributions to mobility engineering, particularly in the area of automotive acoustics and standards. Mr. Hillquist is the founder of the SAE Noise and Vibration Conference which was started in 1985. In 2001 SAE International and the SAE Foundation established an educational scholarship award in his honor.

SAE Fellowship status is the highest grade of membership bestowed by the Society of Automotive Engineers. It recognizes outstanding engineering and scientific accomplishments by an individual that have resulted in mean-

ingful advances in automotive, aerospace, and commercial vehicle technology.

Theodore M. Farabee named a Fellow of ASME

Theodore M. Farabee, ASA Fellow and a mechanical engineer at the Naval Surface Warfare Center, Carderock Division, has been named a Fellow of the American Society of Mechanical Engineers (ASME). Dr. Farabee earned his doctorate in mechanical engineering from The Catholic University of America, Washington, DC.

The fellow grade is conferred upon an ASME member with at least 10 years of active engineering practice and who has made significant contributions to the profession.

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.

2004

24–28 May	75th Anniversary Meeting (147th Meeting) of the Acoustical Society of America, New York, NY [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; E-mail: asa@aip.org; WWW: http://asa.aip.org].
12–14 July	Noise-Con 2004, Baltimore, MD [Institute of Noise Control Engineering of the USA, Inc., INCE/USA Business Office, 212 Marston Hall, Iowa State Univ., Ames, IA 50011-2153; Tel.: 515-294-6142; Fax: 515-294-3528; E-mail: ibo@inceusa.org; WWW: http://www.inceusa.org/NoiseCon04call.pdf]
3–7 Aug.	8th International Conference of Music Perception and Cognition, Evanston, IL [School of Music, Northwestern Univ., Evanston, IL 60201; WWW: www.icmpc.org/conferences.html].
20–24 Sept.	ACTIVE 2004—The 2004 International Symposium on Active Control of Sound and Vibration, Williamsburg, VA (INCE Business Office, Iowa State Univ., 212 Marston Hall, IA 50011-2153; Fax: 515 294 3528; E-mail: ibo@ince.org; WWW: inceusa.org].
15–19 Nov.	148th Meeting of the Acoustical Society of America, San Diego, CA [Acoustical Society of America, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2360; Fax: 516-576-2377; Email: asa@aip.org; WWW: http://asa.aip.org].

2005

16–19 May	Society of Automotive Engineering Noise & Vibration Conference, Traverse City, MI [Patti Kreh, SAE International, 755 W. Big Beaver Rd., Ste. 1600, Troy, MI 48084; Tel.: 248-273-2474; E-mail: pkreh@sae.org].

Cumulative Indexes to the *Journal of the Acoustical Society of America*

Ordering information: Orders must be paid by check or money order in U.S. funds drawn on a U.S. bank or by Mastercard, Visa, or American Express credit cards. Send orders to Circulation and Fulfillment Division, American Institute of Physics, Suite 1NO1, 2 Huntington Quadrangle, Melville, NY 11747-4502; Tel.: 516-576-2270. Non-U.S. orders add $11 per index.

Some indexes are out of print as noted below.

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

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

Volumes 21–30, 1949–1958: JASA, Contemporary Literature, and Patents. Classified by subject and indexed by author and inventor. Pp. 952. Price: ASA members $20; Nonmembers $75.

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

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

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

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

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

Volumes 65–74, 1979–1983: JASA and Patents. Classified by subject and indexed by author and inventor. Pp. 624. Price: ASA members $25 (paperbound); Nonmembers $75 (clothbound).

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

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

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

BOOK REVIEWS

P. L. Marston
Physics Department, Washington State University, Pullman, Washington 99164

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

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

Fundamentals of sound with applications to speech and hearing

William J. Mullin, William J. Gerace, Jose P. Mestre, and Shelley L. Velleman

Allyn and Bacon, 2003
xiii+225 pp. Price: $51.00 (paperback), ISBN: 0-205-3708-7X.

This is one of those books written on a rather elementary level, aimed at a particular target readership. Let me quote from the Preface: ''Students and workers from several fields, such as communication disorders, linguistics, and health sciences, are interested in the mechanisms of speech and hearing. The physics of sound is basic to the understanding of how we communicate, but students in these fields often have not taken a general physics course and may not possess the background for attaining a deep understanding within their own area of study. This book is meant to provide the background in the physics of sound, at a level appropriate for undergraduate non-physical-science majors, for further study of speech and hearing.'' ... ''Our students, while bright and hardworking, often do not have strong backgrounds in science and mathematics. They need material that provides qualitative insight into the nature of sound. We provide that in this book by the frequent use of graphical illustrations. Animations are included on an associated Web site. We use math at the level of algebra because without that, understanding is simply too imprecise to be able to develop the ideas at a level to promote conceptual understanding. However, the math is always accompanied by a thorough verbal interpretation. There is no necessity for the reader to have had previous experience in physics or a physics course.''

Math at the level of algebra! It is true; the authors have in fact kept their promise. Even the Fourier series treatment is done without trigonometry (the use of sines and cosines)!

Let us amble through some of the chapters, making comments that come to mind as we go along. The basic concept of the *wave* receives a very strong treatment in the book, as it should. Chapter 1, ''Introduction to Waves,'' is followed by Chap. 2, ''Standing Waves in Ropes and Strings,'' and Chap. 3, ''Standing Waves in Air Columns.'' In this treatment, the website associated with this book plays an important role. To quote the authors again, ''Sound is a wave, and as such it is a dynamic quantity. Frozen graphics are inadequate to convey understanding without a struggle by the reader. For this reason, the website contains animations that show such situations much more clearly.'' Your reviewer agrees.

Chapter 4 is on the subject of Resonance. This is the place for the reviewer to express his discomfort with the term *resonant frequency*. As the late Frederick V. (Ted) Hunt used to say, the correct usage is *resonance frequency*, because the frequency is not resonant. It is the frequency of resonance!

The subject of waves continues to be treated in Chap. 5, ''Wave Fronts,'' and Chap. 6, ''Complex Waves.'' These chapters introduce practical aspects of wave theory needed for what follows. The reviewer would like to comment on this latter chapter. The authors say, ''...our hearing cannot distinguish the relative phases of the constituents of a complex wave, a rule known as *Ohm's acoustical law*.'' As an electrical engineer, I was initially distressed at this apparent ''hijacking'' of Ohm's law, the basic equation of electricity! Actually, I am now aware that Ohm made contributions in the study of the human auditory system as well.

We go from the time domain to the frequency domain in Chap. 7, ''Wave Analysis.'' This is a good chapter, but I do wonder if the word ''theorem'' in *Fourier's theorem* might not frighten those in the target readership.

The ''meat'' of the book is found in Chap. 8, ''Speech Production,'' and Chap. 12, ''The Ear and the Hearing Mechanism.'' These are very educational for me. The *spectrogram* is introduced at the end of Chap. 8, and plays a strong role in Chap. 9, ''Acoustics of Specific Speech Sounds.'' I like the way the concept and its application are explained.

The next two chapters, Chap. 10, ''Work and Energy,'' and Chap. 11, ''Perception of Sound Intensity,'' will be quite interesting to us acousticians. It is in these chapters that the extremely important, but difficult to explain, concept of the *decibel* is introduced. I am happy to say that the authors have done a very good job with this difficult task.

In summary, despite my minor ''nit-picking,'' I like this work very much. The authors have succeeded in their goal of clarifying what can be an esoteric subject, in my opinion quite admirably. There is a very good index and a very good set of exercises. I definitely recommend the use of this text in courses whose students are in the target audience defined by the authors.

JAMES F. BARTRAM
94 Kane Avenue
Middletown, Rhode Island 02842

OBITUARIES

Edith L. R. Corliss • 1920–2003

Edith L. R. Corliss, a Fellow of this Society, died 4 January 2003 at the University of Connecticut Health Center in Farmington, Connecticut.

Mrs. Corliss was born in Cleveland, OH in 1920. Following her graduation from high school in 1937, she attended the Massachusetts Institute of Technology, where she was awarded Bachelor and Master of Science degrees in Physics. Her education continued over the years with advanced studies at George Washington University, Johns Hopkins University, University of Maryland, American University, and Catholic University of America.

In 1941, Mrs. Corliss joined the United States National Bureau of Standards (NBS, later renamed the National Institute for Science and Technology, NIST); after 1 year she transferred to the Weather Bureau, and then served briefly as an astronomer with the U.S. Naval Observatory. She then returned to NBS in 1944, serving in the Sound Section of the NBS Mechanics Division, continuing this association until her retirement in 1980. In 1963, she was a visiting researcher working with E. Colin Cherry in the Department of Electrical Engineering, Imperial College of Science and Technology, London, England. After her retirement, she served as a guest worker with NIST and also, for a 1-year period, with the National Institutes of Health.

Edith was author of a series of letter circulars, monographs, and consumer's guides on the selection and use of hearing aids. These were among the most widely circulated of any group of NBS publications for many years. Under her guidance, a research effort on the calibration of hearing-related measurement instruments and applications resulted in a program of hearing aid evaluation for the U.S. Veterans Administration (VA), which continues to be a service program administered by NIST. Other activies at NIST included measurements leading to understanding of bone conduction of sound vibration in the head, which assisted in the diagnosis of hearing problems and in the design of hearing aids. She also wrote a series of papers studying the precision of measurement of time-varying signals, and she was coauthor of *NBS Handbook 119, Quieting: A Practical Guide to Noise Control*, a publication that was widely circulated.

Edith Corliss was a Fellow of the Acoustical Society of America, served on the Society's Executive Council, 1973–76, and later served as Vice President Elect 1978–79, and as Vice President 1978–80. She was also chair of the Archives and History Committee (1997–1999) and of the Committee on Meetings (1978–1979). Other professional activities included membership in the Institute of Electrical and Electronics Engineers (IEEE), the American Physical Society, and the Catgut Acoustical Society.

Although her interests extended over a range of technical, scientific, and societal areas, she is probably best known for contributions to speech and hearing-related subjects. In 1957, Mrs. Corliss was awarded the Department of Commerce Silver Medal for Meritorious Service in recognition for her ''very valuable contributions to the science of audiometry and speech analysis and for meritorious authorship.'' She was recognized for her contributions to hearing research by the American Speech-Language-Hearing Association with its Distinguished Service Award in 1978. She worked with various standards writing groups concerned with audiometers, hearing aids, thresholds of hearing, and loudness of sound. For many years, she was deeply involved with Washington D. C. organizations providing outreach to the hearing impaired.

In addition to these primary interests in the field of sound and acoustics, Mrs. Corliss participated in two expeditions that accumulated and reported data on spectra of radiation from solar sources. She served as a member of an IEEE Environmental Quality Committee studying acid rain and air pollution issues.

Mrs. Corliss is survived by a son, Peter, of Hebron, CT, and two grandchildren. She was preceded in death by her husband, Charles H. Corliss, also a retired NBS/NIST research physicist.

Alan B. Coppens • 1936–2003

Alan B. Coppens, a Fellow of the Acoustical Society of America and retired Associate Professor of Physics at the Naval Postgraduate School, Monterey, California, died suddenly on 20 July 2003, at his home in Black Mountain, North Carolina. He was 67 years old. Alan was born in Hollywood, CA, on 26 June 1936. After attending Black Foxe Military Academy in Los Angeles, he went to Cornell University in Ithaca, New York, where he earned a Bachelor of Science Degree in Engineering Physics in 1959. He then continued with graduate work at Brown University in Providence, Rhode Island, earning a Master of Science in Physics in 1962, and a Ph.D. in Physics in 1965. His doctoral dissertation, ''Nonlinear Properties of Metals,'' was under the tutelage of Robert Byer.

After graduation from Brown, he joined the faculty of the Physics Department at the Naval Postgraduate School (NPS), where his dynamic teaching style and sharp wit made him popular with students. At the NPS he primarily taught courses in acoustics and explosions. He advised many Master's theses written by students from both the Engineering Acoustics and the Anti-Submarine Warfare Curricula.

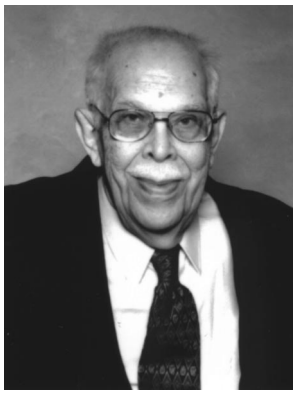
Dr. Coppens is the author of many published papers and monographs on nonlinear acoustics and underwater sound. Of particular importance is his work on nonlinear standing waves in waveguides and the propagation of sound in a wedge-shaped ocean, and for being the first to predict the anomalous disappearance of guided waves during upslope propagation. Among his many publications is the section on ''Sound'' in the *World Book Encyclopedia* and the section on ''Finite-Amplitude Standing Waves'' in the *Encyclopedia of Acoustics*. He is the coauthor of the 3rd and 4th editions of *Fundamentals of Acoustics*, originally published in 1956 under the authorship of Larry Kinsler and Austin Frey of the Naval Postgraduate School, which to this day remains a popular reference and textbook throughout the world.

At the Naval Postgraduate School, he served on many faculty committees and served a term as the Chairman of the Faculty. His research achievements were recognized by his being made a Fellow of the Acoustical Society of America. His activities in the ASA included service on the Technical Committee on Physical Acoustics. He promoted interest in science to young people by presenting courses on the Physics of Music for the Lyceum and serving for many years as a judge for the Monterey County Science Fair.

After his retirement from the Naval Postgraduate School in 1995, Dr. Coppens moved to North Carolina, where he did volunteer teaching in the Physics Department at Warren Wilson College. He is survived by his wife, Linda Miles Coppens, and by his children, Adam B. Coppens and Dale Joan Mason.

JAMES V. SANDERS

William Smith Cramer • (1914–2003)



William Smith Cramer, a Fellow and a former Vice President (1975–1976) of the Acoustical Society of America (ASA), died in Silver Spring, MD, on 27 October 2003. He was a long-time member of ASA and participated with distinction in various functions and activities of the Society, including service on the Executive Council and on the Medals and Awards Committee. He served as Associate Editor of the *Journal of the Acoustical Society*, on one occasion handling papers in Underwater Acoustics and, on another, editing the section on Acoustical News USA. In addition, Dr. Cramer served as editor of the *U.S. Navy Journal of Underwater Acoustics*, JUA (USN), a classified journal, for 23 years. For these services, he was awarded the Distinguished Service Citation of the ASA in 1984.

Dr. Cramer was born in Frederick, MD on 25 August 1914. His early childhood and teen years were spent in Washington, DC, and York, PA. After high school, he attended Ursinus College, in Collegeville, PA, where he received a Bachelor's degree in physics in 1937. From there he went on to Brown University and completed a Master of Science degree (Sc.M.) in 1938. During the 1938–1939 and 1939–1940 academic years, he taught mathematics and science at the University of Maryland, College Park, MD, and at Pikesville College, Pikesville, KY. In the fall of 1940, he returned to Brown University for Ph.D. study, but the completion of his studies was postponed because of the research and development activities associated with the US participation in World War II.

In the spring of 1941, Dr. Cramer took a six-month assignment at the Cruft Laboratory of Harvard University where he worked with a small group, headed by Leo Beranek, on a war-related project to reduce the noise in aircraft. This exposure to research in acoustics and the association with this group of capable men (Rudolph Nichols, Wayne Rudmose, Robert Wallace, and others) had a strong influence on his interest in acoustics. In the fall of 1942 he joined the staff of the Naval Ordnance Laboratory (NOL) in White Oak, MD, and completed his Ph.D. in 1948 while at NOL. His major research effort was the development of methods for studying the physical properties of viscoelastic materials and in the application of these methods in the study of various natural and synthetic rubber compounds. One of the significant developments was the construction of a ''pulse tube'' to measure the reflecting properties and the acoustic impedance of any given anechoic coating. His research in anechoic coating was heavily influenced by the work of Professor Erwin Meyer and his group at the Heinrich-Hertz Institute of the Technical University of Berlin. Meyer and his group (including Professor Eugen Skudrzyk, later at The Pennsylvania State University) were concerned with developing a coating to make their U-boats less visible to sonar and also developing an anechoic lining for water-filled test tanks.

In 1956, Dr. Cramer joined the Office of Naval Research (ONR) in Washington, DC as a scientific staff officer. His main responsibility was to select and monitor a contract program in Physical Acoustics. ONR sponsored a wide range of programs by established researchers (Professor F. V. Hunt, I. Rudnick, Cyril Harris, Uno Ingard, and others) as well by new researchers in the field. He also had the opportunity in his own research activities to make wide contact with almost all workers then active in the field of anechoic coatings. Subsequently, he was assigned the additional duties as Editor of the JUA(USN) in 1959.

In 1966, Dr. Cramer left the Office of Naval Research and the JUA(USN) to become the Head of the Structural Acoustics Branch of the Ship Acoustics Department at the David W. Taylor Naval Ship Research and Development Center. In that capacity, he supervised programs on structural acoustics including some of his own special interest in the properties of material and sound absorbing layers. Later in 1970, he resumed the additional duties of Editorship of the JUA(USN).

Dr. Cramer retired from the Civil Service in June 1976. However, he continued as Editor of the JUA(USN) via an ONR contract with MAR Associates in Rockville, MD, until 30 September 1983. In addition to his editing duties, he served as a consultant on anechoic coatings for industry and the Navy. In all, he wrote and published 46 papers, including a very comprehensive summary of the work that had been done up to that time in the field of acoustic coatings for underwater sound.

In October 1983, the Office of Naval Research nominated Dr. Cramer for the Department of the Navy Meritorious Public Service Citation for his contribution to Naval acoustics as Editor of the Journal. The Chief of Naval Research at the opening of the 35th Naval Symposium on Underwater Acoustics on 1 November 1983 awarded the Citation to Dr. Cramer.

His wife of 55 years, Patricia, their two daughters, Marian and Ann, and their grandchildren, Eric and Maria, survive him.

JOHN T. C. CHEN
R. CRAIG OLSON

REVIEWS OF ACOUSTICAL PATENTS

Lloyd Rice
11222 Flatiron Drive, Lafayette, Colorado 80026

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

Reviewers for this issue:

GEORGE L. AUGSPURGER, *Perception, Incorporated, Box 39536, Los Angeles, California 90039*
ALIREZA DIBAZAR, *Department of BioMed Engineering, University of Southern California, Los Angeles, California 90089*
JOHN ERDREICH, *Ostergaard Acoustical Associates, 200 Executive Drive, West Orange, New Jersey 07052*
MARK KAHRS, *Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261*
HASSAN NAMARVAR, *Department of BioMed Engineering, University of Southern California, Los Angeles, California 90089*
DAVID PREVES, *Starkey Laboratories, 6600 Washington Ave. S., Eden Prarie, Minnesota 55344*
DANIEL R. RAICHEL, *2727 Moore Lane, Fort Collins, Colorado 80526*
WILLIAM THOMPSON, JR., *Pennsylvania State University, University Park, Pennsylvania 16802*
ERIC E. UNGAR, *Acentech, Incorporated, 33 Moulton Street, Cambridge, Massachusetts 02138*
ROBERT C. WAAG, *Univ. of Rochester, Department of Electrical and Computer Engineering, Rochester, New York 14627*

6,659,964

43.20.Ye SONAR-CONTROLLED APPARATUS FOR THE DELIVERY OF ELECTROMAGNETIC RADIATION

Jonathon Lipman, assignor to Neuroscience Tool Works, Incorporated
9 December 2003 (Class 600/557); filed 28 March 2002

A hand-held heat source incorporates a simple sonar system so that the distance from the heat source to the target area, which may be a portion of a human body, a paint-drying surface, a surface etching, etc., can be measured and strength compensations can be made to insure that the incident heat intensity upon that area remains constant.—WT

6,644,118

43.25.Uv CYLINDRICAL ACOUSTIC LEVITATOR/CONCENTRATOR HAVING NON-CIRCULAR CROSS-SECTION

Gregory Kaduchak and Dipen N. Sinha, assignors to The Regents of the University of California
11 November 2003 (Class 73/570.5); filed 13 September 2002

A low-power apparatus for levitation and/or concentration of aerosols and small liquid/solid samples that include particulates is constructed from a hollow piezoelectric crystal. The crystal is formed with a cylindrical cross-section and tuned so that the resonant frequency of its breathing mode matches that of the interior cavity. As the circular cross-section of the transducer is deformed, the resulting acoustic force is concentrated along axial regions parallel to the transducer axis, resulting in concentration of particles in these regions.—EEU

6,652,331

43.30.Sf TROLLING MOTOR WITH INTEGRAL SONAR TRANSDUCER

Robert W. Healey, assignor to Brunswick Corporation
25 November 2003 (Class 440/6); filed 7 July 2001

This patent discusses the concept of mounting the transducer for a fish-finder/depth-finder sonar system within the housing of a small electric trolling motor. Specific attention is directed to the problem of providing improved shielding against sources of electrical noise.—WT

6,661,742

43.30.Sf TROLLING MOTOR WITH SONAR TRANSDUCER

Ronald P. Hansen, assignor to Johnson Outdoors Incorporated
9 December 2003 (Class 367/173); filed 12 October 2001

This patent discusses the concept of mounting the transducer for a fish-finder/depth-finder sonar system within the housing of a small electric trolling motor.—WT

6,645,144

43.35.Sx ELECTROACOUSTIC IMAGING METHODS AND APPARATUS

Han Wen and Robert S. Balaban, assignors to The United States of America as represented by the Department of Health and Human Services
11 November 2003 (Class 600/437); filed 18 May 2001

As indicated in the patent's abstract: ''Methods are disclosed for obtaining electroacoustic images of specimens. One method includes applying an acoustic wave to a specimen and forming an image based on an electroacoustically induced electric field or voltage. In another method, an electric field or voltage is applied to a specimen and an electroacoustic wave is measured to form an image. Apparatus ... is disclosed, as well as methods for distinguishing image contributions from the electroacoustic, thermoacoustic, and Hall effects.''—EEU

6,638,169

43.35.Zc GAMING MACHINES WITH DIRECTED SOUND

Richard L. Wilder *et al.*, assignors to IGT
28 October 2003 (Class 463/35); filed 28 September 2001

The use of demodulated ultrasound is all the rage. Next, consider the cacophony of a gambling casino. The inventors propose using demodulated ultrasound to create a private gaming space. But isn't the sound of gaming machines music to the player's (and casino owner's) ears?—MK

6,640,632

43.35.Zc ULTRASONIC FLAW DETECTION METHOD AND APPARATUS

Hiroaki Hatanaka *et al.*, assignors to Ishikawajima-Harima Heavy Industries Company, Limited
4 November 2003 (Class 73/598); filed 30 April 2002

The arrangement used here is basically the same as that employed in time-of-flight detection methods, but uses wavelet analysis to obtain improved resolution. An ultrasonic transmitter is located on one side of a weld and a receiver on the other side and the two are moved parallel to the weld. The outputs of both are fed via a transceiver to a wavelet analyzer that performs waveform separation by extracting from the waveforms of the flaw detection echoes the highest frequency components that can be extracted using time frequency analysis. The analyzer then extracts subsequent one-half frequency components in sequence. All analyzed components are stored and then multiplied together, and peaks in the results are used to indicate defects. The locations of the defects, identified from the position of the sensors and from the time of flight of the signals, are displayed as *XY* coordinates.—EEU

6,649,069

43.35.Zc ACTIVE ACOUSTIC PIPING

Matthew M. DeAngelis, assignors to BAE Systems Information and Electronic Systems Integration Incorporated
18 November 2003 (Class 210/748); filed 3 December 2002

Piping in which flowing waste water or other fluids is treated contains an inner pipe that has a series of transducers disposed around its external surface. A buffer zone between the inner and outer pipes provides an acoustic impedance mismatch to assist in concentrating the acoustic energy into the flowing material.—EEU

6,652,689

43.35.Zc PROCESS AND DEVICE FOR ULTRASOUND TREATMENT OF A STRIP OF MATERIAL

Klaus Kubik, assignors to Eduard Kusters Maschinenfabrik GmbH
25 November 2003 (Class 156/73.1); filed in Germany 21 March 2001

This is a process, in part covered by German Patent 195 13 246 C2, which is intended for ultrasonic bonding of nonwoven fabrics made of thermoplastic fibers. Ultrasound is used to treat a strip of material that passes through a treatment gap formed by a roller device that includes a rotating roller body and a sonotrode that is directed radially against the working surface of the roller body. In order to avoid any concentricity defect of the roller body that might result from bending, the roller body is directly cooled below its working surface. The roller body features a number of lengthwise bores distributed uniformly over the circumference, preferably in an arc, through which a cooling medium flows fairly uniformly.—DRR

6,663,578

43.38.Ar OPERATOR SUPERVISED TEMPERATURE CONTROL SYSTEM AND METHOD FOR AN ULTRASOUND TRANSDUCER

Michael Peszynski *et al.*, assignors to Koninklijke Philips Electronics N.V.
16 December 2003 (Class 601/2); filed 11 October 2002

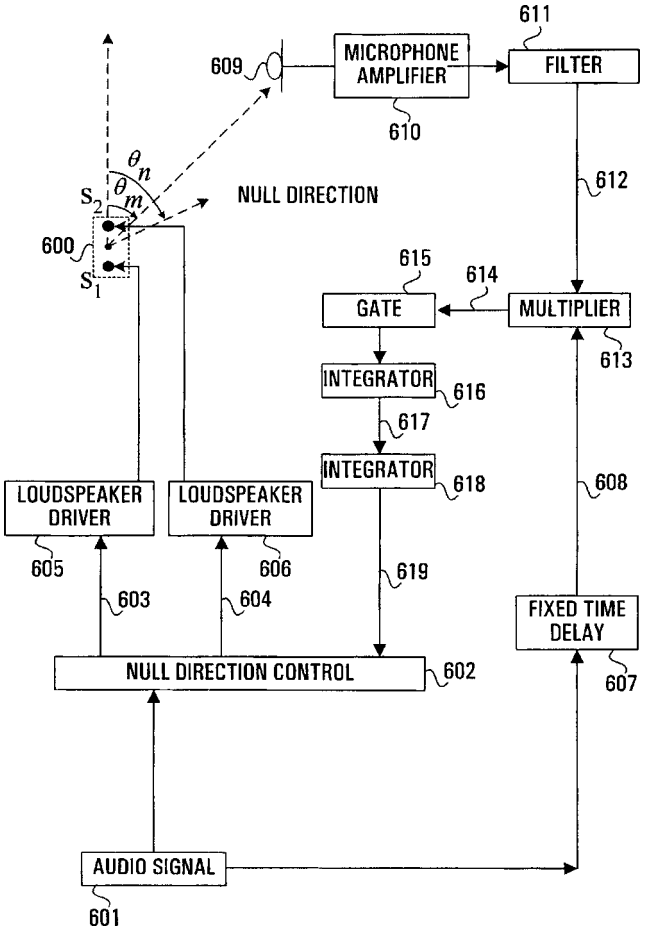
This is a system for cooling an imaging ultrasonic transducer and the patient's contacting surface without the necessity for turning off the transducer. In a preferred embodiment, the operator selects one or more mutable system parameters of the ultrasound system, which will be changed if either the temperature becomes too great, or at the operator's command, or both.—DRR

6,650,758

43.38.Hz ADAPTIVE DUAL PORT LOUDSPEAKER IMPLEMENTATION FOR REDUCING LATERAL TRANSMISSION

Andre J. Van Schyndel, assignor to Nortel Networks Limited
18 November 2003 (Class 381/97); filed 23 December 1999

Suppose that two loudspeakers are mounted on opposite sides of a small enclosure and driven in opposite polarity to operate as a dipole source. By adding variable delay, the pattern can be changed from dipole to cardioid



or anything in between. This kind of steerable null array has already been used to provide privacy at ATM machines and to ''tune out'' acoustic feedback in sound reinforcement systems. The circuitry shown adds adaptive steering to the array.—GLA

6,643,379

43.38.Hz ACOUSTICAL TRANSDUCER FOR RECREATING A SPATIAL SOUND STAGE AND IMPROVED LOCALIZATION OF ORIGINAL SOUNDS SOURCES

Roland Ed Onglao, Cerritos, California
4 November 2003 (Class 381/351); filed 28 March 2000

Fairly conventional vertical arrays are used for left and right stereo loudspeakers. However, the housing for each array is made of flexible material ''...that can vibrate and resonate.'' Moreover, special cutouts are provided on the inner side walls. These techniques are intended to produce skewed coverage patterns similar to those more commonly realized by slightly rotating the speaker cabinets.—GLA

6,656,124

43.38.Hz STACK BASED MULTIDIMENSIONAL ULTRASONIC TRANSDUCER ARRAY

Aimé Flesch and An Nguyen-Dinh, assignors to Vermon
2 December 2003 (Class 600/459); filed 15 October 2001

This design is another variation of transducer arrays based on multilayer piezoelectric elements. The patent also covers a method of manufacturing 1.5D and 2D ultrasonic transducer arrays. The argument is made that the use of piezoelectric multilayer structures as vibrators provides desirable capacitance characteristics and impedance homogeneity. 1.5D transducer arrays are produced so they have configurations having shifted elevation apertures from the center to the edge of the array, which supposedly yields uniformity of transducer capacitance regardless of location within the elevation plane. In production of 2D arrays, dicing a conventional multilayer actuator forms equal-sized transducer elements that are imbedded in a polymer matrix.—DRR

6,659,954

43.38.Hz MICROMACHINED ULTRASOUND TRANSDUCER AND METHOD FOR FABRICATING SAME

Andrew L. Robinson, assignor to Koninklijke Philips Electronics NV
9 December 2003 (Class 600/459); filed 19 December 2001

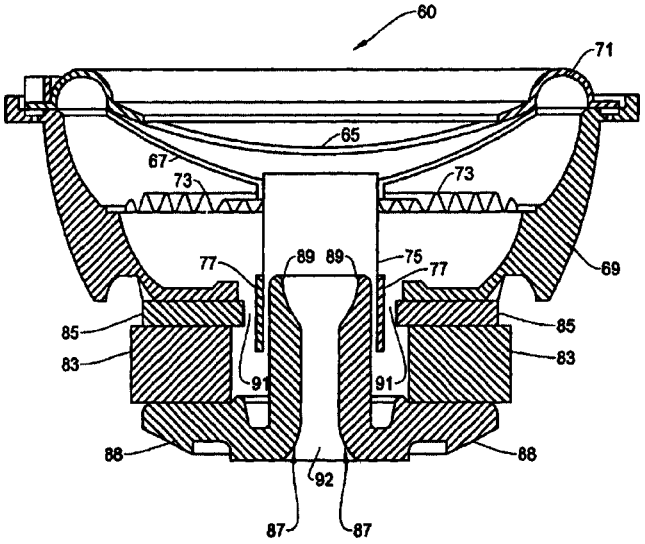
This patent presents improved structures for use with micro-machined ultrasonic transducers (MUTs) and methods for fabricating these structures. In one embodiment, a MUT is formed on a substrate and an acoustic cavity is formed within the substrate at a location below the MUT. The cavity is filled with an attenuation material to absorb acoustic waves propagated into the substrate and to reduce the effect of parasitic capacitances on the operation of the MUT. In a second embodiment, the cavity is formed beneath a group of MUTs and filled with an attenuation material. In a third embodiment, an attenuation material essentially encapsulates a number of MUTs on a dielectric layer. In other embodiments, at least one monolithic semiconductor circuit is formed in the substrate that may be operatively coupled to the MUTs to execute signal processing and/or control operations.—DRR

6,639,993

43.38.Ja LOUDSPEAKER WITH LOW DISTORTION AND HIGH OUTPUT POWER

Jason Kemmerer and Eric Douglas Guarin, assignors to Alpine Electronics, Incorporated
28 October 2003 (Class 381/397); filed 29 December 2001

This woofer uses a long voice coil **77** operating in an unusual magnetic gap. Magnet **83** and outer pole piece **85** are conventional. The inner pole piece takes the form of a straight cylinder extending above and below the



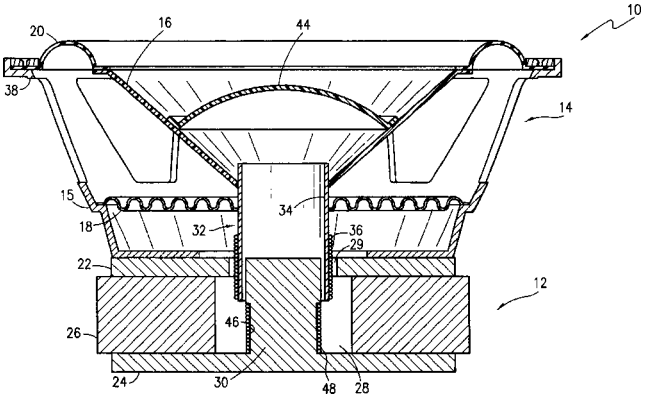
outer pole piece. Moreover, the central vent hole **92** includes a curved flaring section **89** designed to control magnetic flux and produce a symmetrical force versus displacement characteristic.—GLA

6,639,994

43.38.Ja LOUDSPEAKER HAVING ADJUSTABLE MOTOR STRENGTH

Lucio Proni and Brett E. Hanes, assignors to JL Audio, Incorporated
28 October 2003 (Class 381/406); filed 16 August 2000

There was a time when loudspeakers had field coils rather than permanent magnets. Gap flux density could easily be varied over a fairly wide range had anyone wanted to do so. Flux density in a contemporary loudspeaker is usually optimized for a particular application, but electrical system Q can be varied, if desired, by adding a series resistor or adjusting the



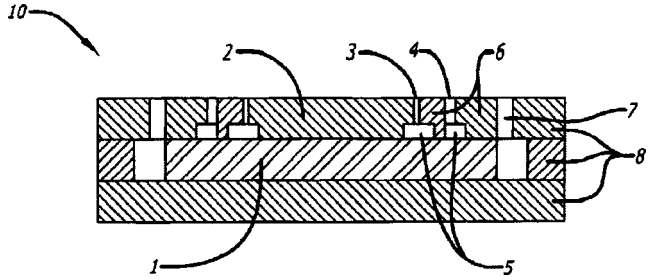
amplifier's internal impedance (variable damping factor). This adjustable loudspeaker has a permanent magnet **26** plus a field coil **46** that can either augment or oppose the flux produced by the magnet. It can certainly be made to work, but consumer demand for such a loudspeaker may be somewhat less than the inventors anticipate.—GLA

6,639,995

43.38.Ja CONCENTRIC MAGNETIC CONFIGURATION FOR LOUDSPEAKERS

You Lian Yu, assignor to Advanced Magnetic Technologies, Incorporated
28 October 2003 (Class 381/412); filed 23 January 2002

The illustration is a section through an interesting multi-gap magnetic circuit for a three-way coaxial loudspeaker. A single magnet **1** provides



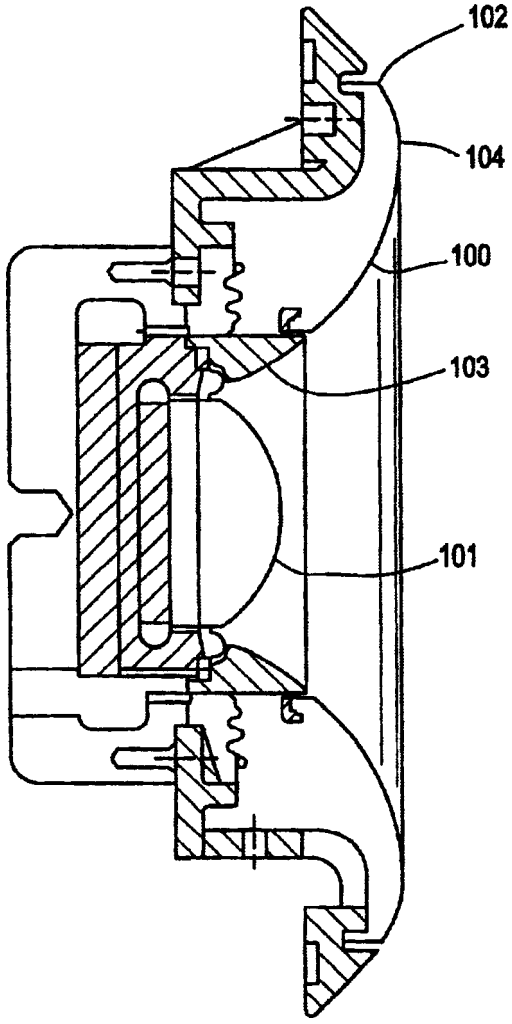
magnetic flux in low-frequency **7**, mid-range **4**, and high-frequency voice coil gaps.—GLA

6,647,122

43.38.Ja LOUDSPEAKER DRIVE UNIT

Andrew Jones, assignor to Pioneer Electronics Technology, Incorporated
11 November 2003 (Class 381/182); filed 16 August 1999

This compound loudspeaker has a high-frequency diaphragm **101**, a short high-frequency horn **103**, and a curved, deformable low-frequency



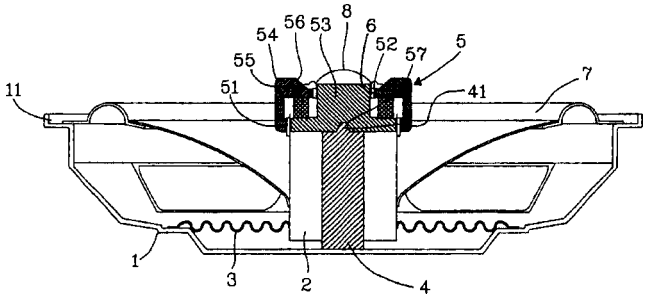
diaphragm **100** fixed at its outer edge **102**. It appears that all three elements must be present to fall under the patent claims.—GLA

6,647,124

43.38.Ja LOUDSPEAKER WITH COAXIAL MAGNETIC CIRCUIT FOR HIGH-PITCH SOUND AND LOW-PITCH SOUND

Jack Peng, assignor to Meiloon Industrial Company, Limited
11 November 2003 (Class 381/412); filed 6 May 2002

This shallow coaxial loudspeaker uses an inverted magnetic assembly to drive the low-frequency cone. Contrary to assertions in the patent, the



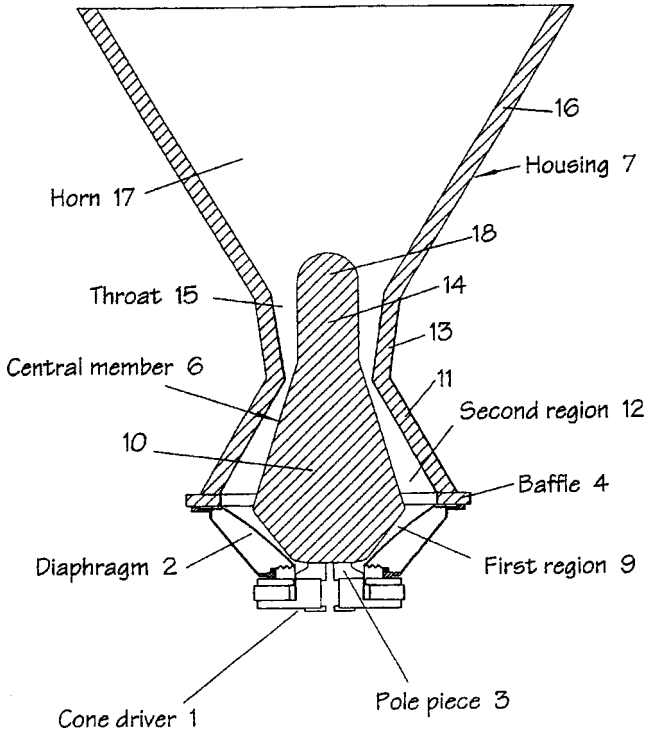
arrangement does not allow high-pitch and low-pitch sounds to be emitted simultaneously.—GLA

6,650,760

43.38.Ja LOUDSPEAKER

Anthony J. Andrews *et al.*, assignors to Funktion One
18 November 2003 (Class 381/343); filed in the United Kingdom 14 July 1999

What you see is a horizontal section through a low-frequency PA horn. Upper and lower surfaces are parallel planes and the horn mouth is rectan-



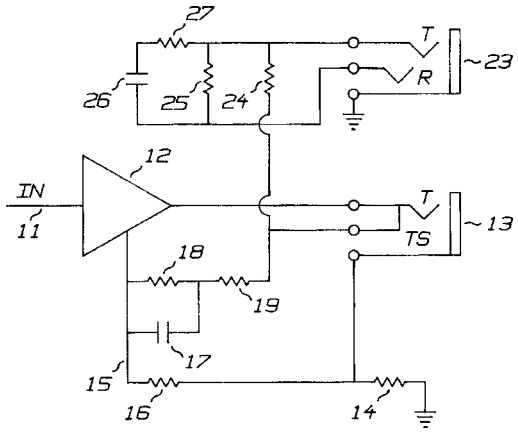
gular. The arrangement is intended to extend the usable range of the horn up to 350 Hz or so.—GLA

6,631,195

43.38.Lc SPEAKER ATTENUATION FOR PRACTICING A MUSICAL INSTRUMENT

Eric K. Pritchard, Berkeley Springs, West Virginia
7 October 2003 (Class 381/120); filed 14 October 1998

Simply, the inventor proposes adding a very simple filter network to



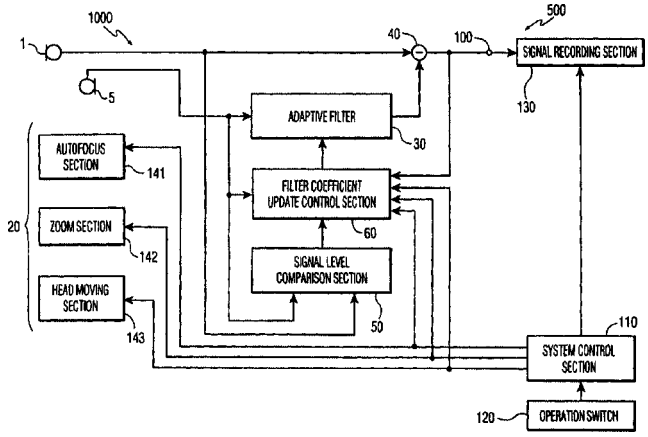
''approximate'' the Fletcher Munson curves for use in practicing or other low level conditions.—MK

6,639,986

43.38.Lc BUILT-IN MICROPHONE DEVICE

Takeo Kanamori *et al.*, assignors to Matsushita Electric Industrial Company, Limited
28 October 2003 (Class 381/71.1); filed in Japan 16 June 1998

It is convenient to have a microphone built into the case of a video recorder so that the user can simply point and shoot. Unfortunately, such a microphone picks up camera noise more efficiently than the desired acoustic source. Earlier patents have disclosed methods of processing the signal from



a second noise-sensing microphone and subtracting it from the main microphone signal. Under conditions of actual use, however, camera noise is variable and unpredictable, making it difficult to update filter coefficients reliably. This patent describes several digital processing techniques intended to minimize echo generation and improve noise cancellation.—GLA

6,650,756

43.38.Lc METHOD AND APPARATUS FOR CHARACTERIZING AUDIO TRANSMITTING SYSTEM, AND METHOD AND APPARATUS FOR SETTING CHARACTERISTICS OF AUDIO FILTER

Nozomu Saito and Tomohiko Ise, assignors to Alpine Electronics, Incorporated
18 November 2003 (Class 381/71.12); filed in Japan 21 May 1997

A method for automatically equalizing an automotive sound system is described. As with other digital equalization schemes, the taps of an adaptive filter are updated to minimize the error between an electronic input signal and the acoustic output signal. ''According to the method for characterizing the audio transmitting system and the method for setting an audio filter, the stable operation of the adaptive filters is ensured without increasing the complexity of the overall system.''—GLA

6,643,372

43.38.Md APPARATUS AND METHOD FOR MUSIC PRODUCTION BY AT LEAST TWO REMOTELY LOCATED MUSIC SOURCES

Dennis L. Ford, Los Angeles, California *et al.*
4 November 2003 (Class 379/388.02); filed 2 January 2003

The use of frequency division multiplexing (FDM) in providing wider bandwidth is well known. Radio stations have been using this technology for at least 10 years. The inventors describe a straightforward two-channel FDM scheme for 3.7-kHz bands. This speedy patent (it only took 10 months!) demonstrates that analog methods live on in the hearts of many. They have conveniently neglected the possibility that two different telephone lines may and can use different trunks, leading to different delays. A digital system could easily calibrate and measure this difference. However, in this system, the inventors do not correct for varying delays of any form.—MK

6,648,031

43.38.Md FUEL DISPENSING DEVICES EQUIPPED WITH A SOUND AND/OR VIDEO SYSTEM

Alan S. Ambrose *et al.*, assignors to Advanced Information Systems, Incorporated
18 November 2003 (Class 141/94); filed 15 August 1995

Pumping gasoline/petrol into a tank is one of life's little pleasures. Imagine how your senses will be enriched by the visual and audible sensations emanating from the boot of the nozzle. Just imagine.—MK

6,650,829

43.38.Md SIGNAL RECORDING APPARATUS AND METHOD

Mariko Kitajima, assignor to Sony Corporation
18 November 2003 (Class 386/120); filed in Japan 22 July 1998

Imagine using different signaling sounds in a video recorder. This is the sole point of this Sony patent.—MK

6,652,346

43.38.Md MUSICAL BOTTLE SYSTEM

Cleveland N. Arnold III, Tampa, Florida
25 November 2003 (Class 446/81); filed 14 June 2002

Yet another design for an electronic musical baby bottle—this time featuring a play button. Think of an MP3 player at the base of a baby bottle and you're on the right track.—MK

6,643,375

43.38.Vk METHOD OF PROCESSING A PLURAL CHANNEL AUDIO SIGNAL

Adam Rupert Philp *et al.*, assignors to Central Research Laboratories Limited
4 November 2003 (Class 381/1); filed in the United Kingdom 25 November 1993

During the past 5 years or so, a lot of engineering time has been devoted to the challenge of producing a convincing 360° sound field from only two loudspeakers. Almost all such efforts involve cancellation of interaural crosstalk by means of electronic filtering to simulate head-related transfer functions (HRTFs). This patent points out that it has not been possible to obtain reliable measurements of HRTFs at frequencies below 200 Hz or so, and that prior art is based on mistaken assumptions when loudspeakers are fairly close to the listener, as in computer games for example. The invention provides a method for calculating and employing an additional ''crosstalk factor'' to produce more convincing sound field simulations.—GLA

6,643,377

43.38.Vk AUDIO OUTPUT SYSTEM AND METHOD THEREFOR

Naoto Takahashi and Hirohiko Katayama, assignors to Canon Kabushiki Kaisha
4 November 2003 (Class 381/306); filed in Japan 28 April 1998

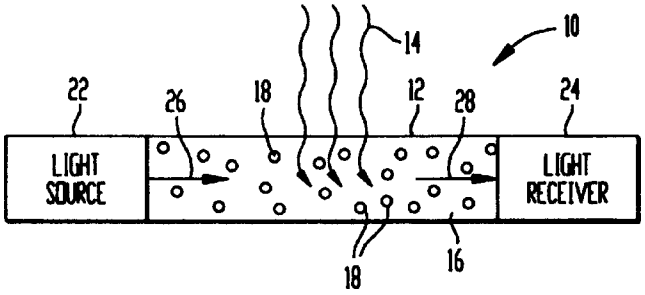
Suppose you are watching a large video image, perhaps a wide screen home theater system or teleconferencing system. The challenge posed by this patent is to make an acoustic source follow the location of a video source and at the same time prevent the sound from being heard by ''irrelevant'' persons. According to the patent, these objects can be attained by ''...a system comprising an image display unit having a screen for displaying the image, a reflective member provided on the screen of the image display unit and adapted to reflect the sound wave, and a speaker for generating a sound wave corresponding to the sound, wherein the speaker is so positioned that the sound wave generated by the speaker is reflected by the reflective member and is emitted toward the side opposed to the screen of the image display unit.'' Fourteen illustrations and four pages of text do not significantly clarify this description.—GLA

6,661,741

43.38.Zp OPTICAL HYDROPHONE AND ARRAY USING BUBBLE RESONANCE FOR DETECTING ACOUSTIC SIGNALS

Thomas R. Stottlemyer, assignor to The United States of America as represented by the Secretary of the Navy
9 December 2003 (Class 367/149); filed 6 June 2002

An optical hydrophone, shown schematically, comprises a housing **12** that is acoustically transparent in some frequency range of interest and that contains a material **16** which is optically transparent but contains a number of air-filled voids **18** such as bubbles or hollow microspheres in suspension throughout the material. This material is typically a viscous liquid or a gel.



When acoustic waves **14** impinge upon the device, the air-filled voids **18** mechanically resonate thereby affecting the transmission of light from source **22** to receiver **24**. The amount of light attenuation can be related to the amplitude of the incident acoustic waves to allow the calibration of the device as a hydrophone. Extension of the concept to a towable linear array of such units, connected in optical cascade fashion, is discussed.—WT

6,641,121

43.40.Tm DAMPING STRUCTURE

Robert P. Carlstedt *et al.*, assignors to Meritor Light Vehicle Technology, LLC
4 November 2003 (Class 267/219); filed 3 January 2002

A perforated steel sheet is mounted between two other steel sheets and a viscous fluid is contained in the spaces between these sheets. When the assembly is deformed, the fluid is forced through the aforementioned perforations, thus resulting in damping.—EEU

6,641,489

43.40.Tm GOLF CLUB SHAFT WITH SUPPRESSED VIBRATION MODES

David E. Wright and Mustapha Hayouna, assignors to Karsten Manufacturing Corporation
4 November 2003 (Class 473/318); filed 26 June 2002

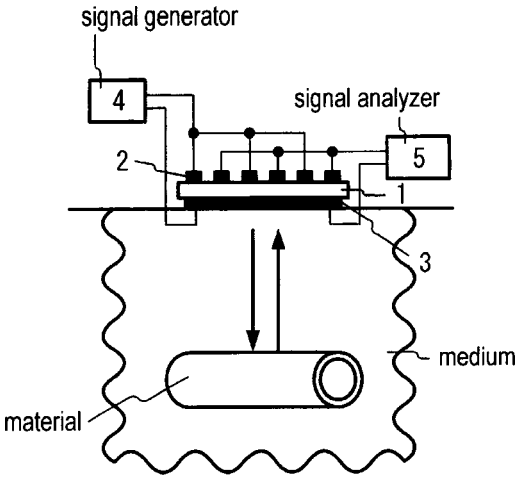
Stiffeners are located along the shaft of a golf club, so as to have the nodes of at least the second and third flexural vibration modes of the club occur at the club face and in the area of the grip. The stiffeners may be made of a shape memory alloy, so that their locations can be changed to suit different users.—EEU

6,637,268

43.40.Yq VIBRATION DISPLACEMENT SENSING SYSTEM

Kohji Toda, Yokosuka, Japan
28 October 2003 (Class 73/602); filed 20 May 2002

An interdigital transducer is made to act in the thickness mode, so as to radiate longitudinal waves into a medium. Two comb-shaped electrodes are applied on the upper surface of a piezolectric substrate and a counter electrode is formed on the lower surface of the substrate. The counter electrode is placed in contact with the medium (such as cellular tissue) in which a vibrating object (perhaps a blood vessel) is to be investigated. Electrical



signals are input between the first comb-shaped electrode and the substrate, and the resulting reflected longitudinal waves are detected between the second comb-shaped electrode and the substrate. The vibration displacement is determined from the interference between the electrically coupled signal between the two electrodes and the detected delayed reflected signal.—EEU

6,640,631

43.40.Yq SYSTEM AND MEASURING SOUND VELOCITY IN MATERIAL

Kohji Toda, Yokosuka, Japan
4 November 2003 (Class 73/597); filed 20 May 2002

The physical configuration described in this patent is identical to that of United States Patent 6,637,268, by the same inventor and reviewed above. Here, electrical signals with given carrier frequencies are applied and the signals resulting from interference between the applied signals and those from the delayed reflections are used to estimate the sound velocity.—EEU

6,640,637

43.40.Yq VIBRATION DETECTING APPARATUS AND METHOD THEREOF

Seung Taek Baek, assignor to LG Electronics Incorporated
4 November 2003 (Class 73/652); filed in the Republic of Korea 20 February 1999

A sensor, particularly for use on drum-type washing machines, consists of a tubular bobbin with a coil wound around its outside and a core free to move inside it. As the core is moved relative to the bobbin due to an applied vibration, the inductance of the coil is monitored via a Colpitts oscillating circuit, which provides information about the vibration.—EEU

6,640,644

43.40.Yq APPARATUS AND METHOD FOR DETECTING TILT AND VIBRATION OF A BODY

Pedro G. Mireles and Jose L. Almaraz, assignors to Delphi Technologies, Incorporated
4 November 2003 (Class 73/721); filed 19 July 2002

This sensor, intended for automotive applications, consists of an edge-supported rectangular piezoresistive membrane that has a weight attached near its center and a piezoresistor near each of its edges. The outputs of the resistors are processed to provide information concerning the membrane's tilt and vibration. The whole assembly is intended to be fabricated from a silicon wafer by use of known photolithographic and etching processes.—EEU

6,661,229

43.50.Gf RF BIRDCAGE COIL WITH REDUCED ACOUSTIC NOISE

Daniel J. Weyers *et al.*, assignors to GE Medical Systems Global Technology Company, LLC
9 December 2003 (Class 324/318); filed 30 April 2001

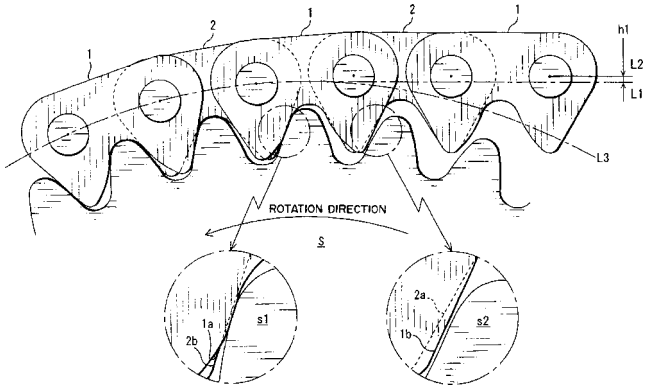
There are a number of acoustical problems associated with the use of magnetic resonance imaging (MRI). The combination of high magnetic fields and gradient switching sequences, for example, results in induced magnetic forces that cause a significant amount of acoustic noise, which can be quite uncomfortable for both the patient and medical personnel. A source of the noise is traceable to the rf coil, typically constructed as a birdcage coil (ref. United States Patent 4,783,641). The birdcage coil consists of ladder-like resonant networks constructed of solid conductors, which are made wide to minimize ohmic resistance in the coil. As the gradient coil is pulsed, a time-varying magnetic field is applied, yielding eddy currents in the conductors. The eddy currents, in turn, induce a perpendicular Lorentz force that causes the coil to vibrate and produce acoustic noise, which becomes increasingly problematic at higher switching speeds. In the ''quiet birdcage'' rf coil described in this patent, a number of slotted, nonconductive channels are incorporated in the conductive material in order to minimize the size of the conductive paths available for eddy currents. Additionally, blocking capacitors are applied between the leg and end segments of the conductive elements to isolate the end segments from the legs at low-frequency harmonics, while effectively acting as short circuits at high operational frequencies.—DRR

6,663,522

43.50.Gf RANDOM ARRANGEMENT TYPE SILENT CHAIN

Hiroshi Horie, assignor to Tsubakimoto Chain Company
16 December 2003 (Class 474/212); filed in Japan 23 February 2001

This random-arrangement-type ''silent'' chain restricts run up and reduces noise and vibration in the free span of the chain. It also limits abnormal elongation and excessive loading during power transmission. One type of link plate begins to engage a sprocket tooth at its inner crotch and sits on the sprocket teeth by its outer flanks. A second type of link plate engages



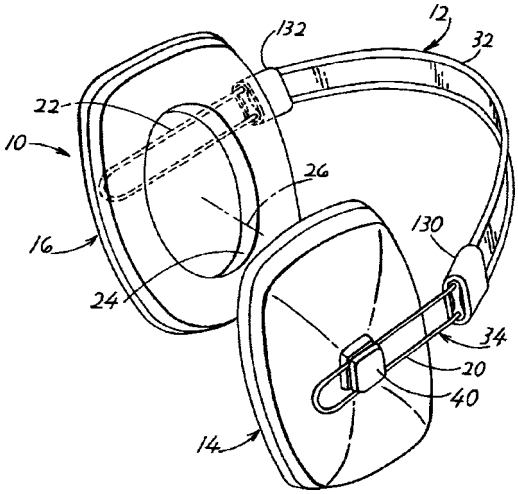
with and sits on the sprocket teeth only by engagement with the outer flanks. These two types of link plates are mixed and combined in the chain. The mixture ratio of the first type of link plate to the total number of link plates is less than 50%. One or more of the second type of link plate are arranged before and after each of the first type, with the number varying randomly along the length of the chain.—DRR

6,611,963

43.50.Hg WIRE BAND EARMUFF

Edwin Woo *et al.*, assignors to Bacon USA Safety, Incorporated
2 September 2003 (Class 2/209); filed 26 October 2001

This is a less expensive circumaural hearing protector fabricated by



using a wire headband **34** covered by a flexible polymer sheath **32**.—JE

6,646,548

43.58.Wc ELECTRONIC SIREN

David P. Dornfeld, assignor to Whelen Engineering Company, Incorporated
11 November 2003 (Class 340/384.4); filed 9 January 2001

Mechanical sirens on emergency vehicles are an endangered species due to the high current drain, component wear, and velocity-dependent behavior. This very modern substitute uses a class-D switcher coupled to a microcontroller. The patent writing is clear and full schematics are provided. Heat up the soldering iron.—MK

6,651,376

43.58.Wc METHOD AND APPARATUS FOR ATTRACTING FISH USING A SOUND PRODUCING FISHING ACCESSORY PROVIDING BLIND POSITIVE ENGAGEMENT TO A HOOK

Donald J. Link, assignor to Skirts Plus Corporation
25 November 2003 (Class 43/42.31); filed 26 June 2002

The author proposes using a two-ball rattle as a fish lure. Many differing embodiments are shown using this basic concept. Caveat piscus.—MK

6,655,165

43.50.Ki AIR CONDITIONER WITH POWER RECOVERY DEVICE HAVING A SOUND SUPPRESSION DEVICE

Ronald S. Eisenhour, assignor to Nissan Technical Center North America, Incorporated
2 December 2003 (Class 62/296); filed 19 December 2002

In an air conditioner, the condenser receives compressed refrigerant from the compressor and condenses it either into a liquid phase or a saturated liquid-vapor phase. The condensed refrigerant passes through a throttle to expand the refrigerant in an evaporator. The refrigerant passing through the throttle is regulated to maintain a regulated refrigerant pressure range within a motor, which gives rise to potentially annoying sounds and vibration. A sound suppression device is introduced between the condenser and the throttle. It consists of a reservoir adapted to hold varying amounts of refrigerant and containing an orifice adapted to damp a pressure wave propagating from a location upstream from the throttle/evaporator system.—DRR

6,653,279

43.64.Gz TREATMENT OF INNER EAR HAIR CELLS

Wei-Qiang Gao, assignor to Genentech, Incorporated
25 November 2003 (Class 514/2); filed 23 August 2000

The patent refers to inducing, promoting, or enhancing the growth, proliferation, or regeneration of inner ear tissue, particularly inner ear epithelial hair cells. It describes methods, compositions, and devices for prophylactic and therapeutic treatment of inner ear disorders and conditions, including hearing impairments. The methods consist of administration of insulin-like growth factor-1 (IGF-1) and/or fibroplast growth factor-2 (FGF-2), or their agonists, alone or in combinations.—DRR

6,658,118

43.50.Ki SUPPRESSION OF FLUID-BORNE NOISE

Satyendra Kumar *et al.*, assignors to Dana Corporation
2 December 2003 (Class 381/71.1); filed 6 December 1999

This noise-canceling apparatus for suppressing fluid-borne noise contains a vibration sensor that provides an electrical signal as a function of the fluid pressure fluctuations in the fluid conduit. The signal is relayed to an electronic controller that responds by energizing a piezoelectric actuator 180° out of phase. The piezoelectric actuator is mounted on the conduit and it imparts out-of-phase pressure fluctuations to the fluid, thereby canceling, at least partially, the noise inside the conduit.—DRR

6,660,757

43.64.Gz USE OF RILUZOLE IN THE TREATING ACOUSTIC TRAUMAS

Jean-Marie Stutzmann and John Randle, assignors to Aventis Pharmaceuticals Incorporated
9 December 2003 (Class 514/367); filed in France 15 December 1998

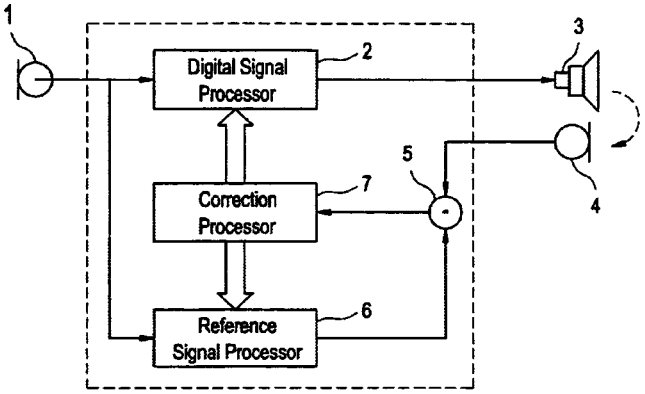
The patent covers the use of riluzole or one of its pharmaceutically acceptable salts in the prevention and/or treatment of acoustic traumas, e.g., deafness and tinnitus. Rizole (2-amino-6-trifluoromethoxy-benzothiazole) is marketed for the treatment of amyotrophic lateral sclerosis and also useful as an anti-convulsant, schizophrenic drug, and in the treatment of Parkinson's disease, mitochondrial diseases, etc.—DRR

6,658,122

43.66.Ts METHOD FOR IN-SITU MEASURING AND IN-SITU CORRECTING OR ADJUSTING A SIGNAL PROCESS IN A HEARING AID WITH A REFERENCE SIGNAL PROCESSOR

Soren Erik Westermann and Morten Kroman, assignors to Widex A/S
2 December 2003 (Class 381/312); filed 6 August 2001

A hearing aid is configured with a second microphone which provides a measure of the output signal delivered. The signal processing of the hearing aid is adjusted in real time by monitoring an error signal derived by comparing the signals from the two microphones.—DAP



6,658,126

43.66.Ts HEARING AID COMPATIBLE PIEZOELECTRIC SPEAKER

Trampas B. Stern, assignor to Ericsson Incorporated
2 December 2003 (Class 381/331); filed 7 April 1999

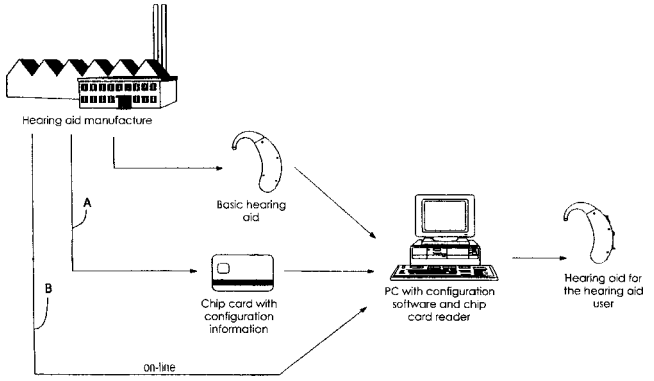
Hearing aids have traditionally sensed signals from land-line telephones via an internal induction coil picking up the magnetic leakage from the handset receiver. Radio telephones commonly use piezoelectric speakers which do not emit significant magnetic leakage. A magnetic element is coupled to the moving portion of the piezoelectric speaker such that a time-varying magnetic field is generated.—DAP

6,658,307

43.66.Ts METHOD FOR CONFIGURING THE FUNCTIONAL PROPERTIES OF AN AUDIOLOGICAL DEVICE

Stefan Mueller, assignor to Siemens Aktiengesellschaft
2 December 2003 (Class 700/87); filed in Germany 14 October 1999

A method is provided that allows a hearing aid dispenser to load firmware into hearing aids prior to initial fitting or to upgrade previously fitted hearing aids. The new information may be transmitted on-line from a manufacturer's data base to a dispenser's programming station or supplied via a separate data carrier such as a chip card. The dispenser may select from different program versions available that are displayed with pricing information.—DAP



6,659,103

43.66.Vt EARPLUG WITH STIFFENER

Jim Tiemens, assignor to Bacou-Dalloz USA Safety, Incorporated
9 December 2003 (Class 128/864); filed 1 August 2002

This is a method of manufacturing a foam earplug with a stiff insert. The insert can accomodate a cord to connect two plugs and facilitates insertion without the need to precompress the foam.—JE

6,654,652

43.66.Yw CALIBRATION AND SECURITY DEVICE FOR PC AUDITORY PROGRAMS

Leslie W. Dalton, Jr. and Robert Brian French, assignors to Beiersdorf AG
25 November 2003 (Class 700/94); filed 23 August 2000

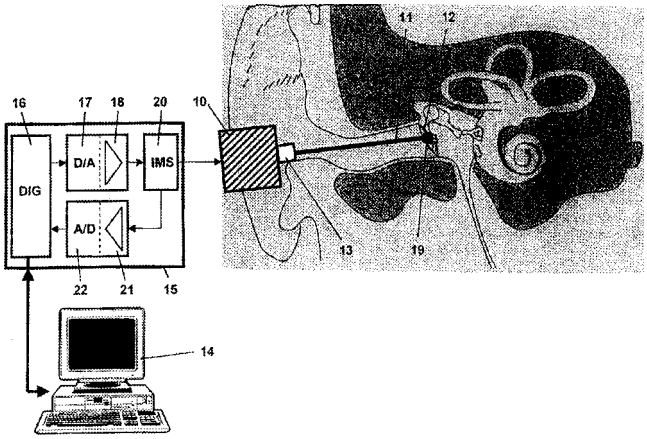
This device is intended to enable a personal computer to perform all the functions of standard audiometric hardware used in the diagnosis and treatment of hearing-related disorders. A typical PC contains sophisticated audio input and output devices, but has no calibration mechanism to regulate the audio output to meet specified sound pressure levels. A hardware module and an installed program initiate a sinusoidal calibration signal on a speaker. As the level of the computer's output signal increases to a reference value, a comparator sends a trigger signal to the PC port. The software directs the computer to increase the output until the desired values corresponding to increased tone volumes are reached, thereby triggering a series of audio signals from the headphones. The interaction of the PC and the hardware module can also act as a theft deterrent for the audiometric software, which is useless without the hardware module.—DRR

6,663,575

43.66.Yw DEVICE FOR ELECTROMECHANICAL STIMULATION AND TESTING OF HEARING

Hans Leysieffer, assignor to Phonak AG
16 December 2003 (Class 600/559); filed in Germany 25 August 2000

The primary object of this hearing testing system is to provide for electromechanical stimulation and testing of hearing in an objective, noninvasive manner. The device consists of an electromechanical transducer for producing mechanical stimulation vibrations, means for noninvasively transmitting the vibrations through the external auditory canal to the (approxi-



mate) center of the tympanic membrane and thereby to the manubrium mallei of the ossicular chain, and an impedance matching system for measuring the mechanical impedance of a biological load structure that is coupled to the transducer.—DRR

6,666,295

43.66.Yw ACOUSTIC RESISTOR FOR HEARING IMPROVEMENT AND AUDIOMETRIC APPLICATIONS, AND METHOD OF MAKING SAME

Mead C. Killion and Andrew J. Haapapuro, assignors to Etymotic Research, Incorporated
23 December 2003 (Class 181/130); filed 23 January 2001

The purpose of acoustic resistance in transducers and sound channels is well known. In the case of a sound tube, for example, a resistance equal to its characteristic impedance will completely damp the length resonances, yielding a smooth frequency response. This patent discloses an acoustic resistor or damper and method of its manufacture. The damper consists of mesh material and attached mounting material. The mounting material defines an open region for the sound to transmit through the mesh and it provides a mounting surface for installing the damper on a surface surrounding the acoustic port or tube. Care is taken to locate the mounting surface on a plane different from the mesh in order to avoid the possibility of applied adhesive touching the mesh material.—DRR

6,523,008

43.72.Ar METHOD AND SYSTEM FOR TRUTH-ENABLING INTERNET COMMUNICATIONS VIA COMPUTER VOICE STRESS ANALYSIS

Adam Avrunin, Atlanta, Georgia and Matthew DelGiorno, Rockville, Maryland
18 February 2003 (Class 704/273); filed 18 February 2000

How can you tell when somebody is lying; especially when the content has been sent to you via the Internet? A number of years ago, there was a flurry of intense interest in using vocal stress as an indicator of truthfulness. The technique turned out to be somewhat less reliable than a polygraph, but not totally without merit. Here, the sender of a statement would add a voice message asserting the statement's truthfulness. A server process would perform a microtremor analysis (not further described in the patent) and transmit the statement along with a report on the veracity of the claim.—DLR

6,519,563

43.72.Fx BACKGROUND MODEL DESIGN FOR FLEXIBLE AND PORTABLE SPEAKER VERIFICATION SYSTEMS

Chin-Hui Lee *et al.*, assignors to Lucent Technologies Incorporated
11 February 2003 (Class 704/246); filed 22 November 1999

Speaker-independent recognition systems are typically more complex than comparable dependent systems, simply because an independent system usually achieves that independence by finding some form of cross-speaker average or composite phonetic representation and, in one way or another, transforming a particular speaker's input into the more general form. This patent basically argues that the simpler form is more appropriate for a speaker recognition or verification task because in that case the speaker idiosyncrasies are just what is of interest. This obvious fact was well known at least 30 years ago.—DLR

6,519,565

43.72.Fx METHOD OF COMPARING UTTERANCES FOR SECURITY CONTROL

Bradley Alan Clements *et al.*, assignors to Voice Security Systems, Incorporated
11 February 2003 (Class 704/273); filed 10 November 2000

To use this speaker verification system, an enrollee must provide a number of example enrollment phrases like those which will be used during system operation. A limited number of points in the time/spectral space are selected for analysis. Within these preselected regions of the spectrogram, the system determines the variability across utterances of amplitude minima and of amplitude values above a preset threshold. In order for an applicant to be accepted, like measurements from the test utterance must fall within the expected range of variability of the enrollment utterances.—DLR

6,529,871

43.72.Fx APPARATUS AND METHOD FOR SPEAKER VERIFICATION/IDENTIFICATION/ CLASSIFICATION EMPLOYING NON-ACOUSTIC AND/OR ACOUSTIC MODELS AND DATABASES

Dimitri Kanevsky and Stephane Herman Maes, assignors to International Business Machines Corporation
4 March 2003 (Class 704/246); filed 25 October 2000

This speaker verification and identification system uses a number of strategies to improve the accuracy of the result. The applicant first speaks certain phrases intended to identify the speaker. The speech is recognized to determine the claimed identity. Information from data records for the purported individual provides a set of queries which are posed to the applicant. The initial speech and the query responses are analyzed for speaker verification. There is only a brief discussion, with little detail, of the method of analysis.—DLR

6,539,352

43.72.Fx SUBWORD-BASED SPEAKER VERIFICATION WITH MULTIPLE-CLASSIFIER SCORE FUSION WEIGHT AND THRESHOLD ADAPTATION

Manish Sharma, Sommerset, New Jersey *et al.*
25 March 2003 (Class 704/249); filed 21 November 1997

This speaker verification system performs a subword analysis using a blind segmentation technique to locate and identify keywords in the candidate speech. By ''blind'' is meant that no linguistic knowledge is used in the

process. This linguistic independence is somehow construed as an advantage for the method. Multiple classification techniques, including channel, fusion, model, and threshold adaptation, as well as dynamic time warping, are used to improve the result.—DLR

6,647,366

43.72.Gy RATE CONTROL STRATEGIES FOR SPEECH AND MUSIC CODING

Tian Wang *et al.*, assignors to Microsoft Corporation
11 November 2003 (Class 704/201); filed 28 December 2001

The inventors describe a multimodal coder such that by using a discrimination algorithm, one can switch from low-bit-rate speech to high-bit-rate audio. The authors are agnostic on the question of discriminator choice, but offer details on the bit-rate management. United States Patent 6,658,383, reviewed below, illustrates how to use such a discriminator in the context of coding mixed content signals.—MK

6,654,716

43.72.Gy PERCEPTUALLY IMPROVED ENHANCEMENT OF ENCODED ACOUSTIC SIGNALS

Stefan Bruhn and Susanne Olvenstam, assignors to Telefonaktiebolaget LM Ericsson
25 November 2003 (Class 703/219); filed in the European Patent Office 20 October 2000

The perceived sound quality of encoded, transmitted, and reconstructed acoustic signals is said to be improved by adding an enhancement spectrum to increase the bandwidth of the system. Coefficients for the enhancement spectrum ranging typically from approximately 4 to 8 kHz are added to the normal primary encoded signals in the 0- to 4-kHz frequency range. The enhancement spectrum contains a larger number of coefficients than the number of sample values in the lower primary spectrum.—DAP

6,654,723

43.72.Gy TRANSMISSION SYSTEM WITH IMPROVED ENCODER AND DECODER THAT PREVENTS MULTIPLE REPRESENTATIONS OF SIGNAL COMPONENTS FROM OCCURRING

Arnoldus Werner Johannes Oomen and Albertus Cornelis Den Brinker, assignors to Koninklijke Philips Electronics N.V.
25 November 2003 (Class 704/500); filed in the European Patent Office 27 August 1999

A method is proposed to prevent multiple representations in two subbands of a single signal component located in the transition bands of the subband filter. In one embodiment, a synthetic audio signal, derived from a longer part of the encoded signal than that representing first signal segments, is subtracted from a signal representing the input signal to form second signal segments. By this process, the first signal segments are removed from the signal, thereby preventing multiple representations.—DAP

6,658,383

43.72.Gy METHOD FOR CODING SPEECH AND MUSIC SIGNALS

Kazuhito Koishida *et al.*, assignors to Microsoft Corporation
2 December 2003 (Class 704/229); filed 26 June 2001

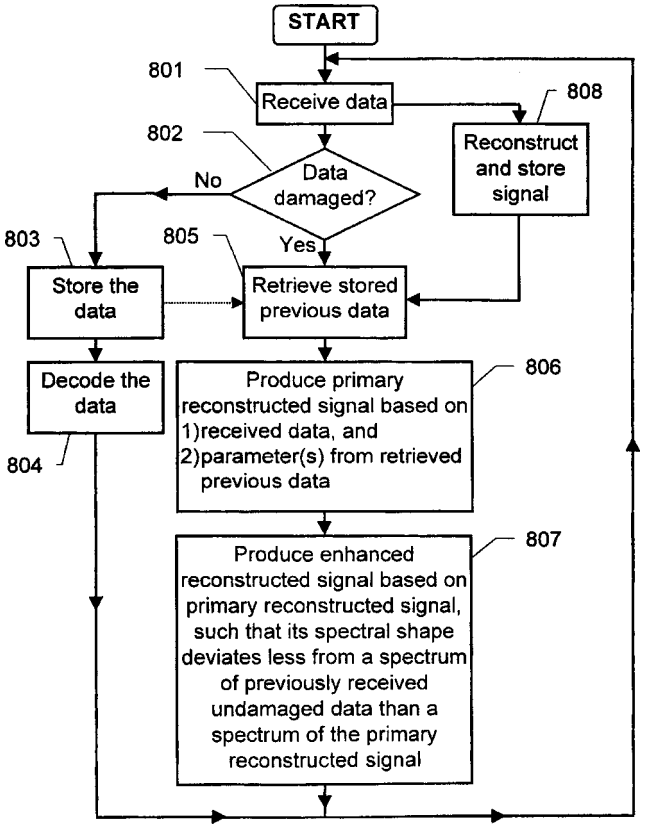
Assigned to Microsoft, this mixed audio/speech coder uses an all-pole model. The program material detector separates speech from audio frames and uses different excitations. In the case of speech, the residual is coded (a la CELP) and transmitted. In the music case, a new overlap-add windowing scheme is introduced before coding. Coding artifacts are no doubt audible for some material, but the authors are silent on this aspect.—MK

6,665,637

43.72.Gy ERROR CONCEALMENT IN RELATION TO DECODING OF ENCODED ACOUSTIC SIGNALS

Stefan Bruhn, assignor to Telefonaktiebolaget LM Ericsson (publ)
16 December 2003 (Class 704/206); filed in the European Patent Office 20 October 2000

Error concealment may be used to improve the sound quality of transmitted acoustic signals if information is lost during transmission. A secondary reconstructed signal is formulated that is based on a primary reconstructed signal and stored data. The method involves utilizing parameters from previously reconstructed, undamaged data stored prior to transmission.—DAP



6,522,248

43.72.Ja MULTICONDITION DETECTION APPARATUS AND METHOD PROVIDING INTERLEAVED TONE AND VERBAL WARNINGS

John J. Andres *et al.*, assignors to Walter Kidde Portable Equipment, Incorporated
18 February 2003 (Class 340/521); filed 18 March 1999

This hazard detector for use in a home or office environment combines the functions of smoke and carbon monoxide detection. In addition to the usual piercing tone, the system then provides a verbal announcement of the nature of the problem. A third set of verbal responses is also available in the case of simultaneous activation of both smoke and CO conditions. The designers anticipate that multiple sets of verbal warnings would be available within the device to provide announcements in alternate languages.—DLR

6,529,874

43.72.Ja CLUSTERED PATTERNS FOR TEXT-TO-SPEECH SYNTHESIS

Takehiko Kagoshima *et al.*, assignors to Kabushiki Kaisha Toshiba
4 March 2003 (Class 704/269); filed in Japan 16 September 1997

What's being clustered here are pitch patterns from normal speech, processed to produce prototype pitch patterns for use by a synthesizer. With each new natural input starting from a random stored pattern, a cluster analysis determines when a new prototype should be created. The final set of patterns is then used for synthesis, each pattern being applied when a similar grammatical structure occurs in the text being synthesized.—DLR

6,535,852

43.72.Ja TRAINING OF TEXT-TO-SPEECH SYSTEMS

Ellen M. Eide, assignor to International Business Machines Corporation
18 March 2003 (Class 704/260); filed 29 March 2001

This patent describes a method of using multiple speakers to produce a database of phonetic patterns for synthesis. The only speech attribute mentioned in the description of the analysis process is the prosodic structure. The single analysis example given is that the running pitch value would be divided by the average pitch, producing a relative-valued pitch curve, which is then taken to be a speaker-independent measure.—DLR

6,615,170

43.72.Ne MODEL-BASED VOICE ACTIVITY DETECTION SYSTEM AND METHOD USING A LOG-LIKELIHOOD RATIO AND PITCH

Fu-Hua Liu and Michael A. Picheny, assignors to International Business Machines Corporation
2 September 2003 (Class 704/233); filed 7 March 2000

This patent proposes a method for voice activity detection which employs frequency features of a windowed speech signal along with pitch information to form Gaussian mixtures. The decision, whether the frame is voiced, is done in two parts. First, a maximum likelihood ratio test is conducted to locally decide between noise and speech and then a tag is attached to the frame. Second, the tags are counted in multiple frames to determine if the input data is speech or noise. The patent argues that the method is robust.—AAD

6,658,385

43.72.Ne METHOD FOR TRANSFORMING HMMS FOR SPEAKER-INDEPENDENT RECOGNITION IN A NOISY ENVIRONMENT

Yifan Gong and John J. Godfrey, assignors to Texas Instruments Incorporated
2 December 2003 (Class 704/244); filed 10 February 2000

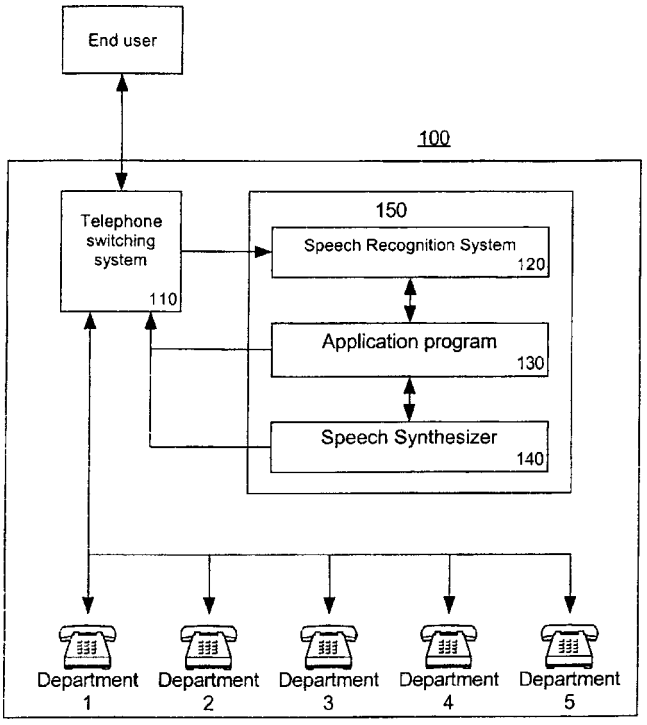
A speaker-independent digit recognizer is trained on a small database using hidden Markov models in a noisy environment.—HHN

6,658,386

43.72.Ne DYNAMICALLY ADJUSTING SPEECH MENU PRESENTATION STYLE

Kimberlee A. Kemble *et al.*, assignors to International Business Machines Corporation
2 December 2003 (Class 704/270); filed 12 December 2000

An adjustable speech menu interface is presented within a system for



routing incoming telephone calls based on an end-user specified call distinction.—HHN

6,662,158

43.72.Ne TEMPORAL PATTERN RECOGNITION METHOD AND APPARATUS UTILIZING SEGMENT AND FRAME-BASED MODELS

Hsiao-Wuen Hon and Kuansan Wang, assignors to Microsoft Corporation
9 December 2003 (Class 704/252); filed 27 April 2000

A speech recognition system uses a combined frame-based and segment-based HMM in order to acquire long-range transition information from the speech signal.—HHN

6,662,159

43.72.Ne RECOGNIZING SPEECH DATA USING A STATE TRANSITION MODEL

Yasuhiro Komori *et al.*, assignors to Canon Kabushiki Kaisha
9 December 2003 (Class 704/255); filed in Japan 1 November 1995

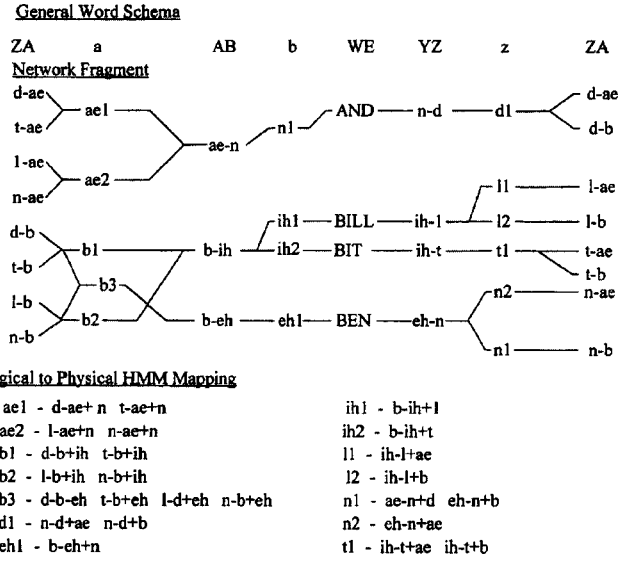
A state transition model is assigned for an unknown word in speech data to reduce the search space and the memory capacity.—HHN

6,668,243

43.72.Ne NETWORK AND LANGUAGE MODELS FOR USE IN A SPEECH RECOGNITION SYSTEM

Julian J. Odell, assignor to Microsoft Corporation
23 December 2003 (Class 704/243); filed in the European Patent Office 25 November 1998

Proposed is a network model structure that utilizes cross-word, context-dependent, triphone hidden Markov models associated with a tree-structured language model that stores all necessary data but is quickly accessible without a large amount of computing power. The language model



Logical to Physical HMM Mapping

ae1 - d-ae+n t-ae+n	ih1 - b-ih+l
ae2 - l-ae+n n-ae+n	ih2 - b-ih+t
b1 - d-b+ih t-b+ih	l1 - ih-l+ae
b2 - l-b+ih n-b+ih	l2 - ih-l+b
b3 - d-b-eh t-b+eh l-d+eh n-b+eh	n1 - ae-n+d eh-n+b
d1 - n-d+ae n-d+b	n2 - eh-n+ae
eh1 - b-eh+n	t1 - ih-t+ae ih-t+b

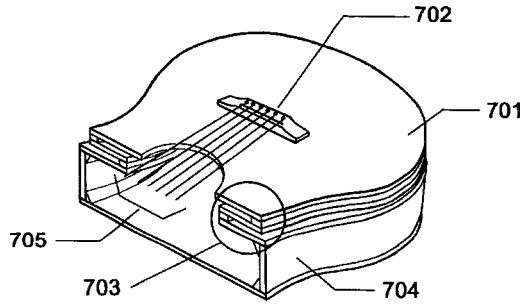
has identifiers associated with each word arranged such that each node of the network model spans a continuous range of identifiers. Tokens transferring through the tree structure may be combined into a set of tokens with similar histories, with likelihood indicating the highest probability of a path representing a word.—DAP

6,646,190

43.75.Gh ACOUSTIC STRINGED INSTRUMENT WITH SPRING SUPPORTED TOP

Kevin Brown, Cambridge, Massachusetts
11 November 2003 (Class 84/291); filed 1 January 2002

The patent teaches that by making the top plate of a guitar more compliant, then the Helmholtz resonance will be broadened with a subse-



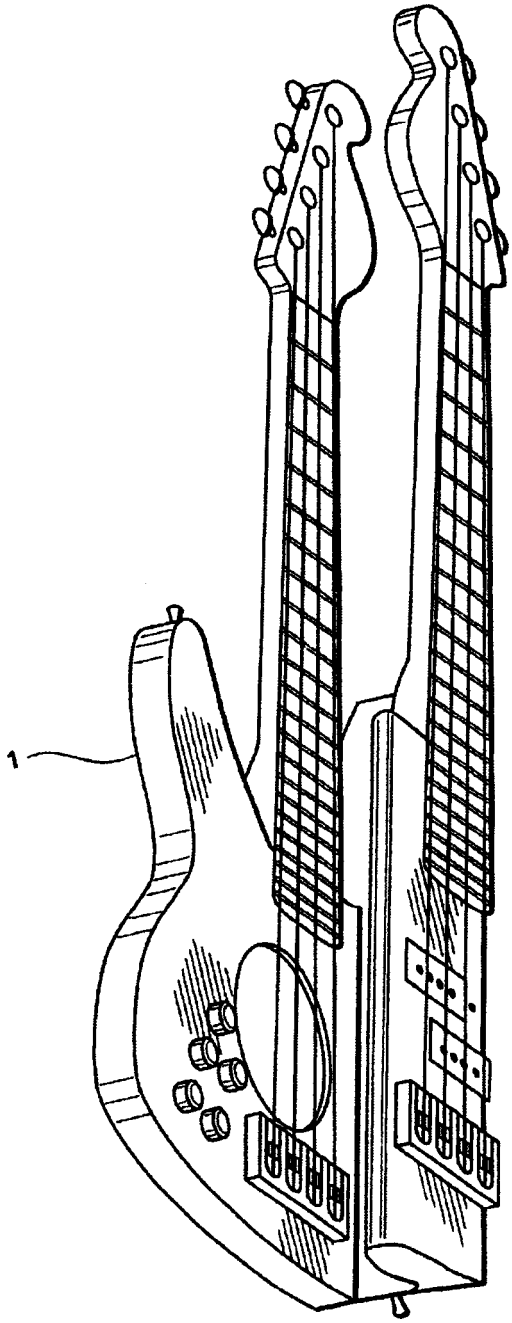
quent improvement in tone. The inventor illustrates a bellows-like attachment scheme as shown in the figure.—MK

6,649,818

43.75.Gh MULTIPLE NECK, INTEGRAL BODY MUSICAL INSTRUMENT

Ronald Irvin Bailey, Tulsa, Oklahoma
18 November 2003 (Class 84/293); filed 26 February 2002

Multiple neck electric guitars are not unknown. Here, the inventor



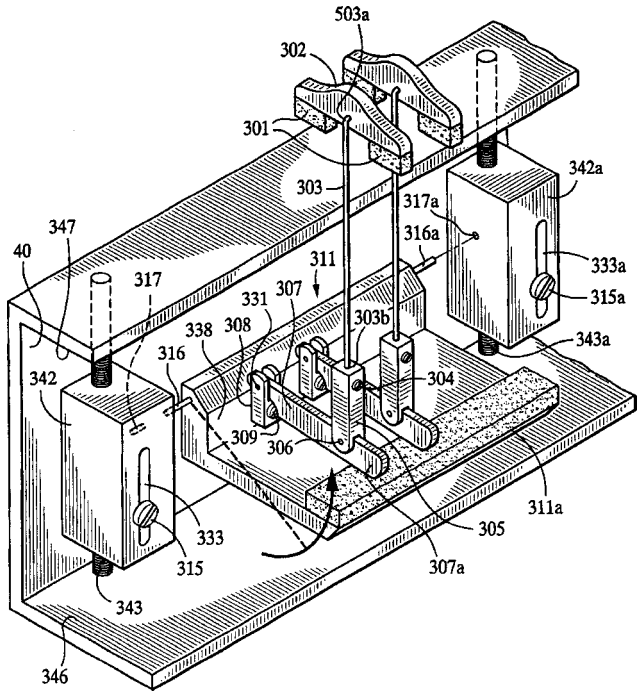
proposes putting the second fretboard literally on top of the other guitar as shown in the figure.—MK

6,632,988

43.75.Mn DAMPER ADJUSTMENT DEVICE

M. Scott Jones and Maik Jenfeldt, assignors to Steinway Musical Instruments, Incorporated
14 October 2003 (Class 84/216); filed 4 April 2000

Steinway describes a new method of damping grand piano strings. The timing of the underdampers creates the action. So, as shown in the figure,



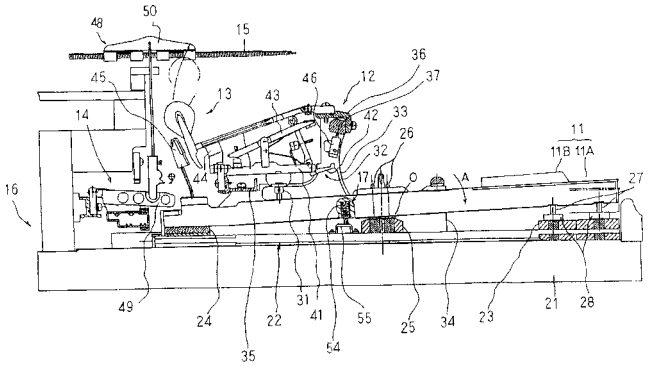
the two adjustment blocks **342** and **342A** control overall position, whereas the individual dampers **301** can be adjusted via the adjustment screw **304**.—MK

6,649,821

43.75.Mn KEYBOARD MUSICAL INSTRUMENT EQUIPPED WITH KEY-TOUCH REGULATOR PROVIDED BETWEEN KEYS AND STATIONERY MEMBER

Satoshi Inoue, assignors to Yamaha Corporation
18 November 2003 (Class 84/423 R); filed in Japan 25 December 2000

Piano action is an individual preference often reinforced by a perform-



er's allegiance to a given manufacturer. Yamaha proposes adding individual adjustments to each key that adjust the time accordingly.—MK

6,653,543

43.75.St MUSICAL INSTRUMENT TUNER WITH CONFIGURABLE DISPLAY

Charles J. Kulas, San Francisco, California
25 November 2003 (Class 84/454); filed 16 May 2002

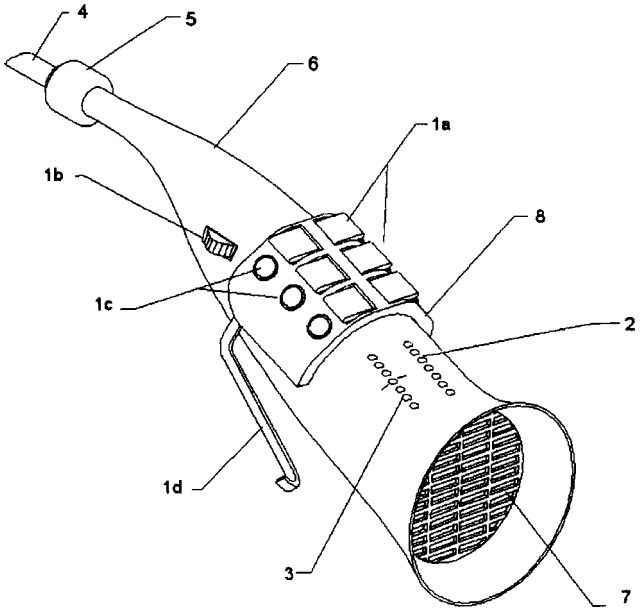
The author proposes a Swiss army knife guitar tuner. First, it can display tunings. Second, it can display gig messages like play lists. Third, it can communicate with anything networked. Lastly, it can be an A/D converter. And... it plugs right into the electric guitar phone jack!—MK

6,653,546

43.75.Tv VOICE-CONTROLLED ELECTRONIC MUSICAL INSTRUMENT

John W. Jameson, assignors to Alto Research, LLC
25 November 2003 (Class 84/735); filed 18 September 2002

A new electronic instrument called the ''Vocollo'' is proposed. The player hums into the mouthpiece. The pitch is detected via the author's own



take on autocorrelation and then used to control a synthesizer that can generate any sound. The addition of keys permits control of various features.—MK

6,647,359

43.75.Wx SYSTEM AND METHOD FOR SYNTHESIZING MUSIC BY SCANNING REAL OR SIMULATED VIBRATING OBJECT

William L. Verplank *et al.*, assignors to Interval Research Corporation
11 November 2003 (Class 703/2); filed 16 July 1999

The authors describe ''scanning synthesis'' whereby some object, any object, can be scanned and converted into music synthesis parameters. Among the more interesting possibilities mentioned is a dancer's body. The key question is how to convert scanned coordinates into control parameters.—MK

6,666,541

43.80.Jz ACOUSTIC EJECTION OF FLUIDS FROM A PLURALITY OF RESERVOIRS

Richard N. Ellson *et al.*, assignors to Picoliter Incorporated
23 December 2003 (Class 347/46); filed 25 September 2001

The device relates to the use of focused acoustic energy for the generation of fluid droplets and their acoustic ejection from multiple fluid reservoirs. The device consists of a number of reservoirs containing the ejection fluid, an ejector incorporating transducers for generating acoustic

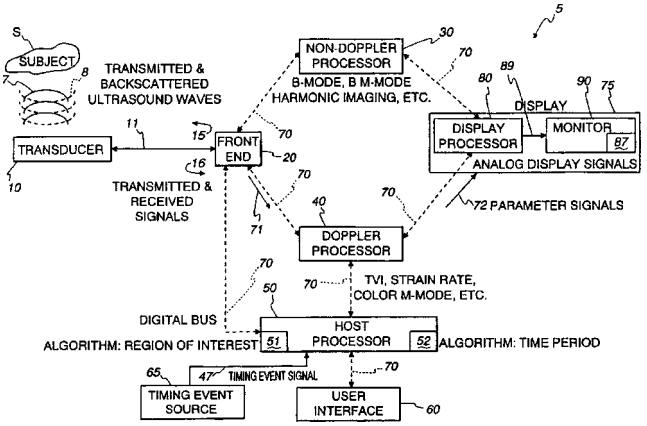
radiation, and a means for focusing the generated acoustic radiation so as to eject fluid droplets from the reservoir fluids. A method is also provided to position the ejector so it can acoustically couple with each of the reservoirs.—DRR

6,652,462

43.80.Qf ULTRASOUND DISPLAY OF MOVEMENT PARAMETER GRADIENTS

Steinar Bjaerum *et al.*, assignors to GE Medical Systems Global Technology Company, LLC
25 November 2003 (Class 600/450); filed 4 June 2002

This device carries the purpose of displaying a color representation of a moving structure, such as a cardiac wall. The color hues correspond to a movement parameter such as velocity or strain rate so that a qualitative visualization is obtained for the movement gradients. The analytic procedure is achieved by transmitting signals into the target structure and then generating received signals in response to ultrasound waves scattered back from



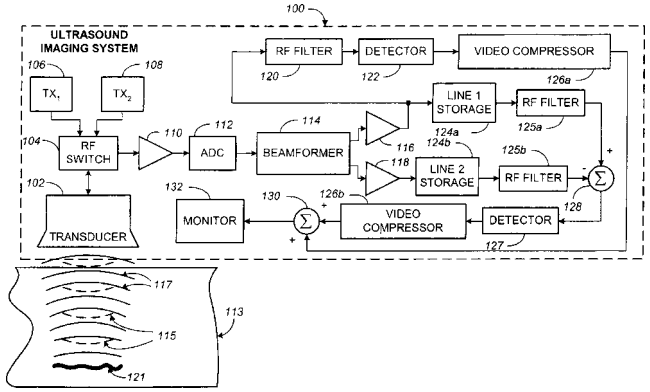
the structure. A processor reacts to the received signals by generating a set of parametric signals representing the values of the movement parameter within the structure during a time interval. A monitor displays the image of the target structure along with color hues corresponding to the plurality of ranges of values of the movement parameter—DRR

6,652,463

43.80.Qf SYSTEM AND METHOD FOR NON-LINEAR DETECTION OF ULTRASONIC CONTRAST AGENTS AT A FUNDAMENTAL FREQUENCY

Thomas J. Hunt *et al.*, assignors to Koninklijke Philips Electronics N.V.
25 November 2003 (Class 600/458); filed 29 October 2002

A system is revealed for real-time imaging of tissue perfused with a contrast agent. This allows for an image with greater nonlinear responses in the perfused tissue that is excited to multiple excitation levels by ultrasound at a fundamental frequency impinging on the tissue. These responses may be gain corrected in an amount corresponding to the difference in magnitude of excitation levels and then mathematically combined to remove linear responses to the fundamental excitation from the obtained image. A transducer



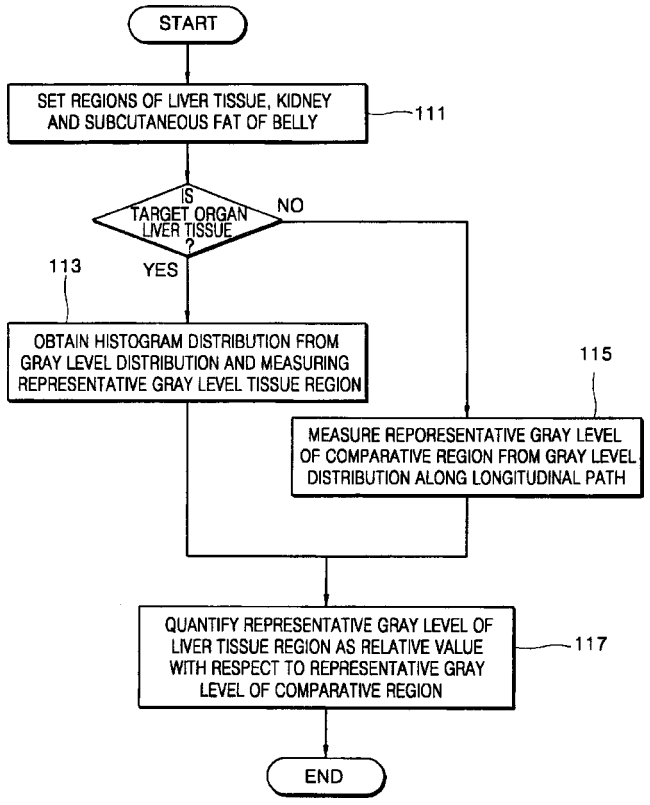
generates two different pressure waves, each being of a different power magnitude. A control system coordinates pressure wave transmissions and reception of ultrasound responses. A processing system removes linear responses, leaving the nonlinear responses from the tissue that is perfused with the contrast agent. An image is produced of the insonified contrast agent and surrounding tissue.—DRR

6,656,121

43.80.Qf METHOD OF QUANTITATIVELY MEASURING FAT CONTENT IN TARGET ORGAN FROM ULTRASOUND VISUAL IMAGE

Ji-wook Jeong *et al.*, assignors to Electronics and Telecommunications Research Institute
2 December 2003 (Class 600/443); filed in the Republic of Korea 31 January 2002

This device uses an ultrasound image to quantitatively assess fat content in a target organ. The procedure includes obtaining an ultrasound visual image of the organ, setting a target region in the obtained image, measuring a quantified representative gray level of the target region from a gray level



distribution of pixels of the target region, and assessing fat content corresponding to the quantified gray level. A typical ultrasound visual image may be an image of a cross section of a belly and the target region may include a human liver.—DRR

6,659,952

43.80.Qf ULTRASONIC DIAGNOSING APPARATUS AND METHOD FOR VIRTUALLY INCREASING POWER OF PIXELS

Yoichi Suzuki, assignors to GE Medical Systems Global Technology Company, LLC
9 December 2003 (Class 600/443); filed in Japan 7 May 2001

In a conventional ultrasonic diagnostic apparatus, adjustments for reducing gain (in the conversion of signal power to pixel brightness) are carried out to inhibit blooming upon the execution of a power display in a Doppler imaging procedure. However, this reduces the capability of projecting thin blood vessels. In this apparatus, designed to prevent the degradation in this capability, when there are two sampling points in a low-power region identical in power but different in variance, the luminance of a pixel corresponding to the sampling point relatively larger in variance is set higher than that of a pixel corresponding to the sampling point with the smaller variance.—DRR

6,661,897

43.80.Qf TRANSDUCER FOR SENSING BODY SOUNDS

Clive Smith, Englewood, Colorado
9 December 2003 (Class 381/67); filed 23 December 2002

This device appears to be some sort of a body microphone adapted to sensing body sounds. The transducer consists of a diaphragm that is placed in contact with a body. The motion of the diaphragm yields a signal representing that motion. This is done by conversion of capacitive, magnetic, or optical conditions to an electrical signal. In this design, the diaphragm would be allowed to move freely without mechanical coupling to a secondary transducer, providing a direct acoustic to electrical conversion means. The transducer would also provide a means for using static diaphragm pressure to control gain and frequency characteristics of the electrical signal.—DRR

6,663,565

43.80.Qf ULTRASOUND DIAGNOSTIC APPARATUS

Tetsuya Kawagishi *et al.*, assignors to Kabushiki Kaisha Toshiba
16 December 2003 (Class 600/437); filed in Japan 3 July 2001

The object of this apparatus is the application of a phase-coded pulse compression scheme for ultrasound diagnosis. A transmitting/receiving unit transmits an ultrasound wave to an object with a specific pulse rate in accordance with a phase code, transmits at a different pulse rate in accordance with a second code complementary to the first code, and receives both reflected signals. The reflected signals are thus each convolved with one of the two codes. The phase difference between the two signals represents the motion of the tissue between the first and second rates. After some compensation of either reception signal on the basis of the phase difference, these two signals are added to generate a third signal, which is used to produce an image.—DRR

6,663,568

43.80.Qf ULTRASOUND TECHNIQUES

Robert Wyatt Gill, assignor to Commonwealth Scientific and Industrial Research Organisation
16 December 2003 (Class 600/456); filed in Australia 11 March 1998

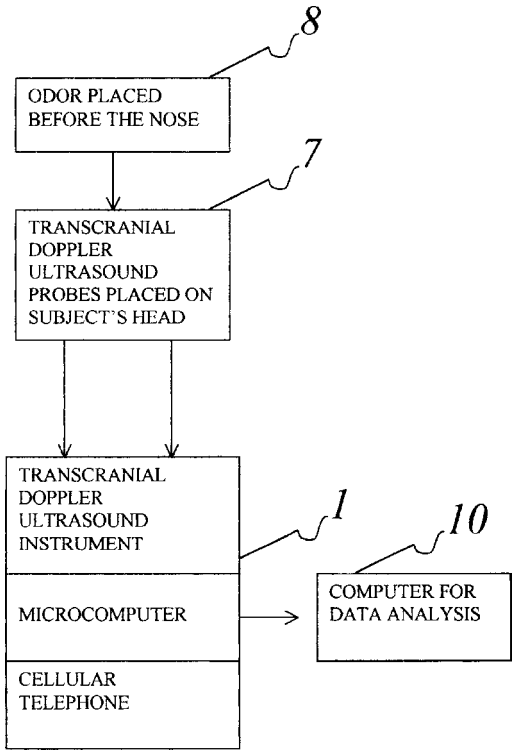
This method estimates volume flow in a vessel through the use of ultrasound techniques that include the steps of imaging the vessel and using an ultrasonic transducer array so as to yield fluid velocity information for the vessel. The plane formed by the transducer beams has a direction offset from the axis of the vessel in such a manner that it forms an elliptical section within the vessel. The velocity information for points within the vessel and the elliptical section are used to calculate a mean velocity for the fluid flow through the elliptical section. The calculation uses a power-weighted average of the velocity information, the power weighting being correlated to the intensity of the received signal at each point corresponding to the velocity information. The assertion is made that this method can be used on the internal portion of human or animal anatomy and can be implemented on a standard ultrasound machine by reprogramming the machine.—DRR

6,663,571

43.80.Qf TRANSCRANIAL DOPPLER ULTRASOUND DEVICE FOR ODOR EVALUATION

Philip Chidi Njemanze, Owerri, Nigeria
16 December 2003 (Class 600/504); filed 28 May 2002

This system for odor matching and odor selection is based on testing a living subject (a human or a scent dog) by presenting the subject with one or more odors under different conditions while monitoring cerebral blood flow velocity of the subject. The system consists of an odor administration container, a transcranial Doppler ultrasound device connected to a microcomputer, a cellular telephone, and a main-frame analysis computer. The device uses a conventional transcranial Doppler instrument with a pulsed wave 2-MHz ultrasonic transducer probe. The probes are positioned on the acoustic window of the temporal bone above the zygomatic arc on both sides of the head and are focused on major cerebral vessels. The device is designed to allow canine and human subjects to move freely in following an odor trail



by integrating a microcomputer and cellular telephone circuitry with the conventional circuitry of the transcranial Doppler system to provide wireless communication with the mainframe computer either directly or via the Internet in real time. It is apparent that some expertise would be required in the evaluation of the Doppler blood flow signals to gauge reaction to odor trails.—DRR

6,652,473

43.80.Sh ULTRASONIC AND GROWTH FACTOR BONE-THERAPY: APPARATUS AND METHOD

Jonathan J. Kaufman and Alessandro Chiabrera, assignors to Jonathan J. Kaufman
25 November 2003 (Class 601/1); filed 21 March 2001

This apparatus pertains to therapeutically treating bone and musculoskeletal tissue *in vivo*, involving both biophysical and biochemical stimuli. The bone is subjected to ultrasound emitted by a transducer placed on the skin of a bony member. The ultrasound is a repetitive finite duration signal consisting of multiple frequencies that are in the ultrasonic range to 20 MHz. A bone growth factor (a biochemical compound that stimulates bone growth) is applied to the adjacent skin prior to stimulation with ultrasound. The ultrasound serves to transport the bone growth factor to the bone and then to synergistically enhance the interaction with the bone, thereby inducing healing, growth, and ingrowth responses. In another embodiment, a vibrational or mechanical input in combination with a biochemical compound may serve to hasten bone fracture healing and to treat osteoporosis.—DRR

6,652,547

43.80.Sh APPARATUS AND METHOD OF REMOVING OCCLUSIONS USING ULTRASONIC MEDICAL DEVICE OPERATING IN A TRANSVERSE MODE

Robert Rabiner and Bradley A. Hare, assignors to OmniSonics Medical Technologies, Incorporated
25 November 2003 (Class 606/159); filed 2 February 2001

A transverse mode ultrasonic probe is introduced into a blood vessel in the proximity of an occlusion and ultrasonic energy is applied to the probe until the occlusion is removed. The probe contains a small cross-sectional lumen and is articulated for navigating a path along a zigzag vessel. The probe can be combined with acoustic and/or aspiration sheaths to promote destruction and removal of an occlusion. It may also be used with a balloon catheter.—DRR

6,656,136

43.80.Sh USE OF FOCUSED ULTRASOUND FOR VASCULAR SEALING

Lee Weng *et al.*, assignors to Therus Corporation
2 December 2003 (Class 601/2); filed 25 October 2000

A method and apparatus are described for sealing vascular punctures and wounds by delivery of ultrasound energy to a vascular site to arrest bleeding. The site of a puncture or wound in the vascular vessel is determined and an ultrasonic transducer is positioned in the vicinity. The transducer applicator is coupled to a control that administers focused ultrasonic energy of sufficient energy and duration to denature tissues at the puncture, thereby closing and sealing the puncture.—DRR

6,660,013

43.80.Sh APPARATUS FOR REMOVING PLAQUE FROM BLOOD VESSELS USING ULTRASONIC ENERGY

Robert A. Rabiner and Brad A. Hare, assignors to OmniSonics Medical Technologies, Incorporated
9 December 2003 (Class 606/128); filed 5 October 2001

This relates to a method and apparatus for removing plaque, including fatty deposits and other occlusions, from blood vessels by ablation through the use of ultrasonic energy. The apparatus features a small cross-sectional profile for removing plaque buildup while avoiding direct contact with the wall and thereby avoiding any potential damage to the artery. It is said that the apparatus can be used in both surgical and outpatient treatment with minimal post-operative complications. The ultrasonic probe may include aspiration channels on its outer surface or an aspiration sheath may enclose the ultrasonic probe so that the location of the aspiration port may be varied axially relative to the ultrasonic tip.—DRR

6,663,554

43.80.Sh ULTRASONIC METHOD AND DEVICE FOR WOUND TREATMENT

Eilaz Babaev, assignor to Advanced Medical Applications, Incorporated
16 December 2003 (Class 600/2); filed 7 August 2002

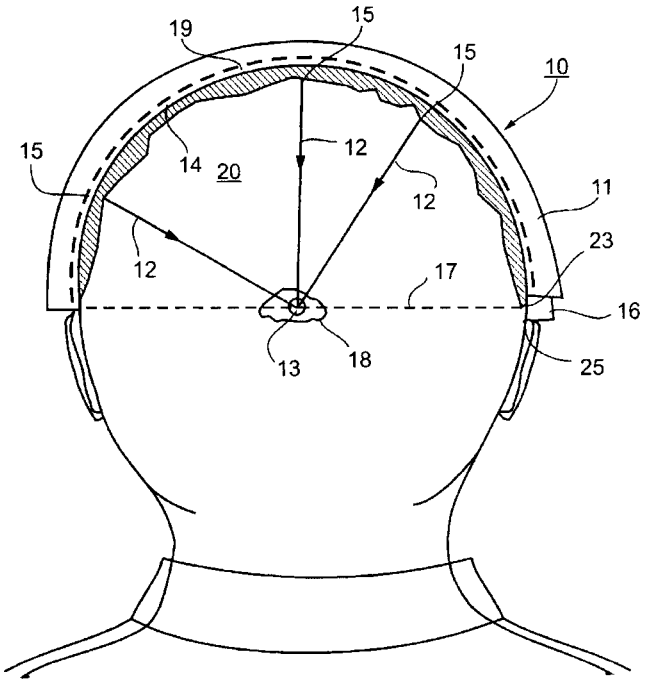
This is another device that uses ultrasound to treat wounds. The ultrasonic transducer serves as a radiating surface from a tip at the device's distal end. A liquid is directed to the radiation surface through a central orifice or a separate tube, resulting in a directed atomized particle spray at the tip. The spray helps to transmit ultrasound waves and provides an irrigation, mechanical cleansing, liquid energizing, and bactericidal effect.—DRR

6,666,833

43.80.Sh SYSTEMS AND METHODS FOR FOCUSSING AN ACOUSTIC ENERGY BEAM TRANSMITTED THROUGH NON-UNIFORM TISSUE MEDIUM

Zvi Friedman and Dov Maor, assignors to Insightec-TxSonics Limited
23 December 2003 (Class 601/2); filed 28 November 2000

This is a thermal energy treatment system that uses a focused acoustic energy beam to transmit through nonuniform tissue. The system features an array of transducer elements located outside of the body. The ultrasound energy is sent into a nonuniform tissue medium, such as a skull, and focused on a location (where a tumor may be situated) inside the body. This procedure is executed by monitoring a signal representing the reflected portions of the incident beam to a desired focal point or zone through a portion of the medium that does not significantly distort the reflected signal. The system



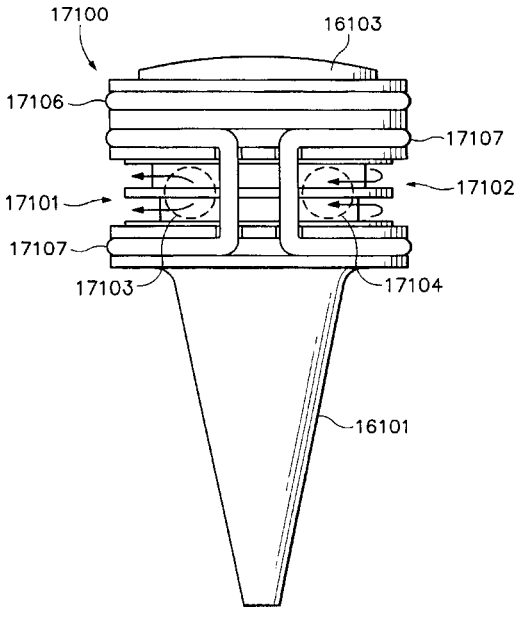
also includes an ultrasound detector having a fixed position relative to the transmitting elements and configured to receive a signal constituting a reflected absorption of the beam along a line-of-sight axis. A controller is coupled to the transmitter and detector, configured to calculate the intensity of the axially reflected portion of the beam based on the timing of the received signal. The controller then adjusts the transmission of one or more of the transmitting elements on the basis of calculated intensity of the reflection portion of the beam along the axis.—DRR

6,666,835

43.80.Sh SELF-COOLED ULTRASONIC APPLICATOR FOR MEDICAL APPLICATIONS

Roy W. Martin and Andrew Hollis Proctor, assignors to University of Washington
23 December 2003 (Class 601/2); filed 8 March 2002

This embodiment of an ultrasonic applicator provides a method for producing a therapeutic, high-intensity, focused, ultrasonic energy pattern in living tissue, while providing active cooling of the transducer without interference of the sonic wave. The applicator utilizes solid tapered cones



mounted to a preferably concave, spherically curved, piezoelectric ultrasound transducer that concentrates the ultrasound energy to a narrow tip so that very high levels of ultrasound can be delivered to the target tissue abutting the tip. The housing of the applicator may include a channel for carrying a coolant to absorb heat generated by the transducer.—DRR

6,652,460

43.80.Vj METHOD FOR ULTRASONIC IMAGING, PARTICULARLY OF MOVING BODIES, SUCH AS SURGICAL UTENSILS, TISSUES, FLOWS, OR THE LIKE

Paolo Pellegretti *et al.*, assignors to Esaote, S.p.A.
25 November 2003 (Class 600/443); filed 4 May 2001

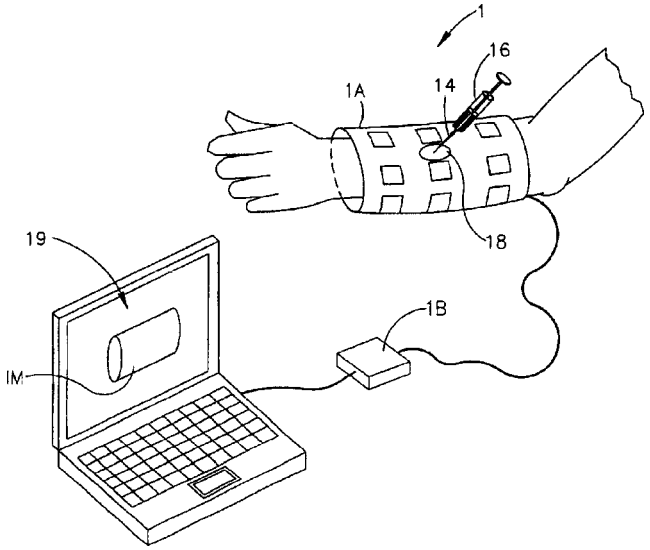
The difference between echoes produced by two successive transmissions is processed nonlinearly to obtain an image in which intraoperative surgical apparatus or structures such as calcification as well as blood flow are shown.—RCW

6,652,461

43.80.Vj ULTRASOUND DEVICE FOR THREE-DIMENSIONAL IMAGING OF INTERNAL STRUCTURE OF A BODY PART

Zeev Levkovitz, assignor to F.R.A.Y. Project Limited
25 November 2003 (Class 600/443); filed in Israel 15 April 1999

This is a portable device that consists of a notebook computer, a control unit, and a flexible cuff containing ultrasonic transceivers. The transceivers are arranged in parallel groups that are spaced apart. The space



between adjacent transceivers in a group results in overlapping regions of ultrasonic illumination. Cross-sectional images are acquired and combined to show three-dimensional structure.—RCW

6,656,120

43.80.Vj ULTRASOUND IMAGING SYSTEM USING KNOWLEDGE-BASED IMAGE ADJUSTING DEVICE

Seong Woo Lee *et al.*, assignors to Madison Company, Limited
2 December 2003 (Class 600/437); filed in the Republic of Korea 16 November 2001

Stored patient information is used to simplify instrument operation. This is accomplished by computing a so-called similarity between reference

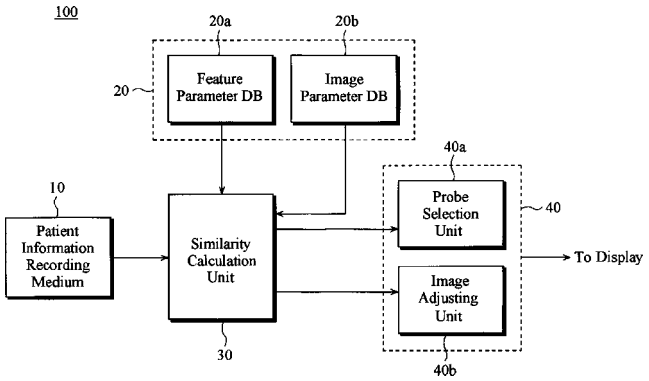


image parameters and patient information. Parameters with the highest similarity are used to select a type of probe and to adjust imaging system settings.—RCW

6,656,123

43.80.Vj COMBINED FUNDAMENTAL AND HARMONIC ULTRASONIC IMAGING AT LOW MI OR DEEPER DEPTHS

Seth E. Jensen and Michalakis Averkiou, assignors to Koninklijke Philips Electronics N.V.
2 December 2003 (Class 600/458); filed 19 December 2001

Pulses are transmitted with two fundamental frequency components. When a contrast agent is present, one of these components produces harmonic components at the fundamental frequency of the other signal and these components are used to image the agent. When a contrast agent is not present, fundamental echoes of the other frequency are used for imaging.—RCW

6,659,953

43.80.Vj MORPHING DIAGNOSTIC ULTRASOUND IMAGES FOR PERFUSION ASSESSMENT

Thilaka S. Sumanaweera and Robert W. Steins, assignors to Acuson Corporation
9 December 2003 (Class 600/443); filed 20 September 2002

Images in a sequence are mapped to one frame of reference by using local warping. During the warping, nonlinear transformations are used to obtain interpolated images in the common reference. The warping produces a sequence of images in which the same spatial location in each image represents essentially the same spatial location of the tissue being imaged.—RCW

6,659,955

43.80.Vj MEDICAL DIAGNOSTIC ULTRASOUND IMAGING SYSTEM TRANSMITTER CONTROL IN A MODULAR TRANSDUCER SYSTEM

Vaughn R. Marian, Jr., assignor to Acuson Corporation
9 December 2003 (Class 600/459); filed 27 June 2002

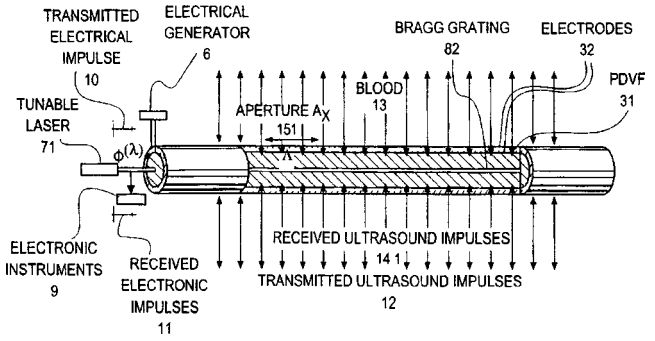
When contacts in a scan head are not electrically coupled with contacts in a receptacle of a modular transducer system, transmitters in an ultrasonic imaging instrument are disabled to prevent high voltages from being present in the receptacle.—RCW

6,659,957

43.80.Vj OPTICAL-ACOUSTIC IMAGING DEVICE

Gil M. Vardi, Town & Country, Missouri and Victor Spivak, late of Kiriat Bialik, Israel
9 December 2003 (Class 600/467); filed 11 June 2002

This device consists of a single mode optical fiber with a Bragg grating and a piezoelectric jacket to image over 360°. A wire in the device can be



used as a guide for vascular interventions such as inflation of a balloon or placement of a stent.—RCW

6,663,566

43.80.Vj METHOD AND APPARATUS FOR AUTOMATIC CONTROL OF SPECTRAL DOPPLER IMAGING

Lihong Pan and Richard Kulakowski, assignors to GE Medical Systems Global Technology Company, LLC
16 December 2003 (Class 600/454); filed 19 February 2002

Lines of Doppler shift are calculated. A processor determines the presence of aliasing and estimates noise levels from the lines. System parameters such as pulse repetition frequency, baseline shift, and spectrum orientation are then automatically adjusted. The processor also determines positive and negative signal boundaries for each spectral line. The boundaries are used to display a spectral trace that corresponds to the edges of the spectral lines.—RCW

6,663,567

43.80.Vj SYSTEM AND METHOD FOR POST-PROCESSING ULTRASOUND COLOR DOPPLER IMAGING

Ting-Lan Ji and Glen McLaughlin, assignors to Zonare Medical Systems, Incorporated
16 December 2003 (Class 600/455); filed 19 March 2002

Doppler parameters are compared to threshold values and masks are produced. This is accomplished using a spatial filter and a classification operator coupled to the spatial filter. Pixels with values not close to values of nearby pixels are reclassified based on the values of neighboring pixels.—RCW

6,666,823

43.80.Vj BEAM COMBINATION METHOD AND SYSTEM

Lin Xin Yao, assignor to Siemens Medical Solutions USA, Incorporated
23 December 2003 (Class 600/443); filed 26 March 2002

Receive beamforming uses combinations of transmit beams to enable ultrasonic imaging at faster frame rates that do not have image degradation normally associated with parallel beamformation and that do not require intensive interpolation.—RCW

6,666,824

43.80.Vj SYSTEM AND METHOD OF DYNAMIC AUTOMATIC SENSING OF AVAILABLE DYNAMIC RANGE

David Rust and David Roundhill, assignors to Koninklijke Philips Electronics N.V.
23 December 2003 (Class 600/443); filed 1 April 2002

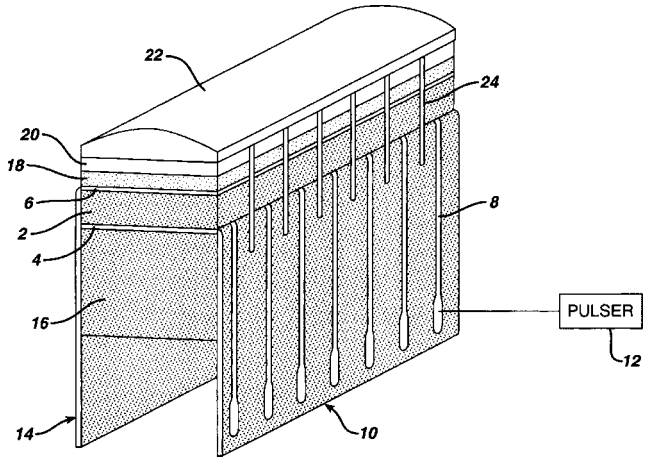
Echoes are processed to determine the noise present. The noise determination is used to create a signature and the signature is used to maximize dynamic range in the presence of noise. The dynamic range and noise rejection are varied for different regions of an image as a function of the scanning geometry being used.—RCW

6,666,825

43.80.Vj ULTRASOUND TRANSDUCER FOR IMPROVING RESOLUTION IN IMAGING SYSTEM

Lowell Scott Smith *et al.*, assignors to General Electric Company
23 December 2003 (Class 600/459); filed 5 July 2001

A silicon acoustic matching layer **18** is next to the piezoelectric layer **2** in this transducer to achieve improved resolution. The silicon layer is part of



an acoustic stack that includes other matching layer materials **20** and ground electrode **6**. The thickness of the silicon layer is nominally a quarter of the wavelength in silicon.—RCW

Effects of pulse rate on threshold and dynamic range in Clarion cochlear-implant users[a] (L)

Heather A. Kreft,[b] Gail S. Donaldson, and David A. Nelson
Clinical Psychoacoustics Laboratory, University of Minnesota, Department of Otolaryngology, MMC 396, Rm 8-323 PWB, 420 Delaware St. SE, Minneapolis, Minnesota 55455

(Received 10 December 2002; revised 30 January 2004; accepted 5 February 2004)

The effects of pulse rate on absolute threshold (THS), maximum acceptable loudness (MAL), and dynamic range (DR) were evaluated in 15 Clarion cochlear implant users. A wider range of pulse rates was assessed than in previous studies, and subjects with both standard and perimodiolar electrode arrays were tested. THS and MAL decreased with pulse rate, and DR increased with pulse rate, for pulse rates between 200 and 6500 pulses per second (pps). However, slopes of THS-vs-pulse rate and MAL-vs-pulse rate functions became shallower above 3250 pps. Subjects with standard electrode arrays had similar THSs as subjects with perimodiolar electrode arrays at all pulse rates. In contrast, subjects with standard arrays had significantly higher MALs and larger DRs than subjects with perimodiolar arrays, and these differences became larger with increasing pulse rate. [DOI: 10.1121/1.1701895]

PACS numbers: 43.66.Cb, 43.66.Ts [MRL] Pages: 1885–1888

I. INTRODUCTION

Cochlear implants have undergone several important changes over the past decade, one of which is the use of increasingly faster rates of pulsatile stimulation: Whereas early devices used pulse rates in the range of 250 pps/channel, contemporary devices are capable of stimulating at rates greater than 5000 pps/channel (Advanced Bionics, 2003). Previous studies have shown that increasing pulse rate reduces electrical threshold and increases electrical dynamic range in both human cochlear implant listeners (Simmons, 1966; Shannon, 1985; 1989; Vandali *et al.*, 2000; Skinner *et al.*, 2000) and animal models (Pfingst *et al.*, 1980; Black *et al.*, 1983; Pfingst *et al.*, 1993; 1995; Miller *et al.*, 1997). In general, these studies have indicated that electrical thresholds decrease -2 to -4 dB per doubling of pulse rate and that most comfortable loudness levels (MCLs) or maximum acceptable loudness levels (MALs) decrease more gradually (-0.8 to -1.2 dB/doubling of pulse rate).[1] However, previous studies were limited to pulse rates of 3000 pps and only a few subjects were tested at rates above 2400 pps. Thus, it is not known whether threshold continues to decrease and dynamic range continues to increase with pulse rate over the entire range of pulse rates available in newer devices.

The design of the implanted electrode array is also known to affect electrical thresholds and dynamic ranges. At low pulse rates ($<$1000 pps), the perimodiolar arrays used in contemporary cochlear implants require substantially less current to generate comfortable (or maximal) loudness levels than the standard arrays used in earlier devices (Osberger *et al.*, 1999; Young and Grohne, 2001; Donaldson *et al.*, 2001; Lesinski-Shiedat *et al.*, 2000; Lenarz *et al.*, 2001; Xu and Pfingst, 2002; Parkinson *et al.*, 2002; and Hay-McCutcheon *et al.*, 2002). There is much less difference in current levels needed to reach detection threshold; as a result, dynamic ranges are smaller for perimodiolar electrode arrays than for standard arrays. Because comparisons between perimodiolar and standard electrode arrays have not been made at pulse rates above 1000 pps, it is not known whether differences in their operating ranges are maintained at the higher pulse rates now in common clinical usage.

The present study was designed to extend previous results by evaluating electrical thresholds and dynamic ranges at higher pulse rates. The first goal was to determine whether the effects of increasing pulse rate observed for pulse rates less than 3000 pps (decreases in threshold and increases in dynamic range) were also observable at higher pulse rates (up to 6500 pps). The second goal was to determine whether previously observed differences in the current requirements of standard versus perimodiolar electrode arrays were constant across a wide range of pulse rates.

II. METHODS

Subjects were 15 postlingually deafened adults with a Clarion version 1.2 cochlear implant. Eight subjects had a standard electrode array (Spiral electrode array [**SPRL**]) and seven subjects had a perimodiolar electrode array (HiFocus electrode array with electrode positioning system [**HF+EPS**]). Subjects had used their devices for 5 months to 4 years prior to participating in the study. Although average duration of implant use was significantly longer for SPRL subjects (2.5 years) than for HF+EPS subjects (1.0 years), average word recognition scores were similar for the two groups (NU-6, SPRL=51.7%, HF+EPS=48.3%). Three individual electrodes distributed across the electrode array were evaluated in each subject (one each from the apical, middle, and basal third of the array).

Experiments were controlled by a personal computer (PC) running the Clarion SCLIN '98 FOR WINDOWS software (Advanced Bionics Corporation, 1996–97), and a second PC running custom software that controlled a research interface

[a]Portions of these data were presented at the 2001 Conference on Implantable Auditory Prostheses (Pacific Grove, CA).
[b]Electronic mail: plumx002@umn.edu

provided by Advanced Bionics Corporation for the Clarion C-I intracochlear stimulator (ICS). Stimuli were 200-ms pulse trains comprised of 77-μs/phase, cathodic-first, biphasic pulses with no interphase gap. Pulse rates were 200, 500, 1000, 1625, 2600, 3250, and 6500 pps. Electrode coupling was monopolar, with the active, intracochlear electrode referenced to a return electrode on the case of the ICS. Nominal current amplitudes specified in the clinical and research stimulation software were translated to calibrated amplitudes using a set of tables developed in our laboratory. These tables compensated for nonlinearities in the current source that were functions of electrode impedance and pulse rate. Electrical impedances for the three test electrodes were measured at the beginning and end of each data collection session using the SCLIN software. Impedance measures typically varied less than $\pm 10\%$ across sessions. The average value of impedance measures obtained at all test sessions was used to compute the calibrated amplitudes for a given electrode.

Data collection for each test electrode spanned two, 2-h research sessions. The middle electrode was tested first, followed by the apical and basal electrodes. In each session, THS and MAL were measured using the SCLIN software with an ascending method of adjustment procedure. Additional threshold estimates were subsequently obtained with a three-interval, forced choice adaptive procedure that used a three-down, one-up stepping rule to estimate the current amplitude corresponding to 79.4%-correct detection (Levitt, 1971). The three listening intervals for each trial were cued visually on a video monitor, and the signal was presented in one interval, selected at random. The subject used a computer mouse to indicate the interval thought to contain the sound. Correct-answer feedback was given immediately after each response. Current amplitudes corresponding to the final eight reversals of each adaptive track were averaged to obtain a single threshold estimate. Adaptive thresholds were obtained in sets where a given set included one threshold estimate at each pulse rate, collected in increasing order of pulse rate. Final values for THS and MAL at each pulse rate were computed by averaging the combined data obtained in the two research sessions (4–6 adaptive threshold estimates and 2 MAL estimates). Dynamic range (DR) was computed by subtracting the final THS from the final MAL at each pulse rate.

III. RESULTS AND DISCUSSION

The left panel of Fig. 1 shows the mean data for THS and MAL as a function of pulse rate for subjects with the SPRL and HF+EPS electrode arrays. Although some individual subjects showed differences in THS or MAL across test electrodes (apical, middle, basal), there were no systematic effects of electrode site on average THS or MAL for either group. Thus, the mean data shown in Fig. 1 were obtained by first collapsing the data across test electrodes for each subject, and then averaging the data across subjects in each group. Error bars in the figure represent standard errors of the means.

It is apparent from Fig. 1 that mean THSs for SPRL and HF+EPS subjects are similar at all pulse rates. This result is consistent with findings of Lesinski-Shiedat *et al.* (2000) and Lenarz *et al.* (2001), who compared standard and perimodiolar electrode arrays for the Clarion device at fixed pulse rates. Other studies conducted at fixed pulse rates have shown slightly lower thresholds for perimodiolar arrays than for standard arrays (Osberger *et al.*, 1999; Young and Grohne, 2001; Donaldson *et al.*, 2001; Xu and Pfingst, 2002; Parkinson *et al.*, 2002; Saunders *et al.*, 2002).

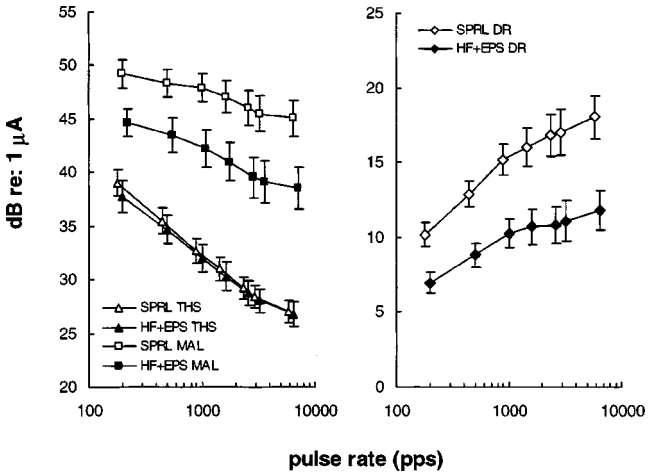


FIG. 1. Left panel: Mean detection threshold (THS) and maximum acceptable loudness level (MAL) as a function of pulse rate, for Clarion subjects with standard (SPRL) and perimodiolar (HF+EPS) electrode arrays. Right panel: Mean dynamic range (DR) for the same groups. Error bars represent ± 1 s.e. of the mean.

Linear functions provided good fits to individual THS-vs-pulse rate functions in log–log coordinates (dB re: 1 μA vs log pulse rate). Slopes of THS-vs-pulse rate functions were nearly invariant across subjects in both the SPRL and HF+EPS groups, with a mean value of -2.4 dB/doubling. This slope is within the range (-2- to -4-dB/doubling) of slopes reported in earlier studies (Simmons, 1966; Shannon, 1985; 1989; Vandali *et al.*, 2000; Skinner *et al.*, 2000).

Figure 1 illustrates that mean MALs were 4 to 7 dB higher for subjects with the SPRL array than for subjects with the HF+EPS array. This result is also similar to data previously reported at fixed pulse rates (Lesinski-Shiedat *et al.*, 2000; Lenarz *et al.*, 2001). A two-way (group$\times$pulse rate) repeated-measures ANOVA confirmed a significant main effect of group [$F(1,78)=6.20$, $p<0.05$] and also indicated that group differences in MAL increased slightly with pulse rate [$F(6,78)=3.92$, $p<0.01$, for the interaction component]. *Post hoc* tests showed that the group differences in MAL were statistically significant for pulse rates of 1000 pps and higher (Tukey test, $p<0.05$), and approached significance at 200 and 500 pps (Tukey test, $p=0.093$ and $p=0.067$, respectively).

Individual MAL-vs-pulse rate functions were also well fit by linear functions in log–log coordinates. Even though mean MALs were considerably different for SPRL and HF+EPS subjects, the slopes of MAL-vs-pulse rate functions were similar across groups. The slopes of these functions (mean$=-1.2$ dB/doubling) were substantially shallower than the slopes of THS-vs-pulse rate functions, and were similar to corresponding slopes for MAL and MCL indicated by the data of Skinner *et al.* (2000) and Vandali *et al.* (2000), respectively.

Although THS-vs-pulse rate and MAL-vs-pulse rate functions were fit with a single linear function between 200 and 6500 pps, some individual functions of both types showed a reduction in slope between 3250 and 6500 pps. These reductions in slope are also evident in the mean data (Fig. 1). Thus, although the effects of pulse rate observed at lower rates appear to be operative at pulse rates as high as 6500 pps in some subjects, there is also some evidence of saturation above 3250 pps. Evaluation of pulse rates higher than 6500 pps would be necessary to determine the pulse rates at which complete saturation occurs.

The right panel of Fig. 1 shows the average DR as a function of pulse rate for SPRL and HF+EPS subjects. Higher MALs led to substantially larger DRs for SPRL subjects (mean=15.2 dB across pulse rates) as compared to HF +EPS subjects (mean=9.3 dB across pulse rates). A two-way repeated-measures ANOVA confirmed that the main effect of group was significant [$F(1,78)=14.80$, $p<0.05$] and a significant group×pulse rate interaction showed that differences between groups increased with pulse rate [$F(6,78)=4.62$, $p<0.001$]. *Post hoc* tests indicated that group differences in DR were significant at all pulse rates (Tukey test, $p<0.05$).

IV. CONCLUSIONS

(1) The effects of pulse rate on THS, MAL, and DR are qualitatively similar at pulse rates above 3000 pps as at lower pulse rates: both THS and MAL improve with pulse rate. However, average rates of improvement are reduced above 3250 pps, suggesting a partial saturation of the underlying mechanisms at very high pulse rates. Further evaluation of pulse rate effects above 6500 pps is needed to determine the pulse rates at which complete saturation occurs.
(2) Clarion cochlear implant subjects with standard (SPRL) and perimodiolar (HF+EPS) electrode arrays have similar thresholds over a wide range of pulse rates (200–6500 pps).
(3) Clarion subjects with standard (SPRL) electrode arrays have significantly higher MALs than those with perimodiolar (HF+EPS) electrode arrays, and these differences increase with increasing pulse rate. As a result, differences in DR also increase with pulse rate.
(4) Slopes of MAL-vs-pulse rate functions are similar for Clarion subjects with standard (SPRL) and perimodiolar (HF+EPS) electrode arrays, even though average MALs are significantly different between groups.

ACKNOWLEDGMENTS

This research was supported by NIDCD Grant DC00110 and the Lions 5M International Hearing Foundation. The authors thank Advanced Bionics Corporation for providing the Clarion research interface; Eric Javel for software development; and Suzanne Hansel and Shanna Allen for assistance in data collection. They also extend special thanks to the subjects who participated in this study. Dr. Marjorie Leek and two anonymous reviewers provided valuable suggestions on an earlier version of the manuscript.

[1]Only two studies reported data for MCL or MAL as a function of pulse rate: We estimated a slope of −1.2 dB/doubling from the MAL data of Skinner *et al.* (2000) for pulse rates between 600 and 2400 pps, and estimated a slope of −0.8 dB/doubling from the MCL data of Vandali *et al.* (2000) for pulse rates between 250 and 1615 pps.

Advanced Bionics Corporation (**2003**). ''New methodology for fitting cochlear implants,'' HiResolution Sound Technology Workshops, Valencia, CA.
Advanced Bionics Corporation (**1996–97**). SCLIN '98 FOR WINDOWS [cochlear implant programming software]. Sylmar, CA.
Black, R. C., Steel, A. C., and Clark, G. M. (**1983**). ''Amplitude and pulse rate difference limens for electrical stimulation of the cochlea following graded degeneration of the auditory nerve,'' Acta Oto-Laryngol. **95**, 27–33.
Donaldson, G. S., Peters, M. D., Ellis, M. R., Friedman, B. J., Levine, S. C., and Rimell, F. L. (**2001**). ''Effects of the Clarion Electrode Positioning System on auditory thresholds and comfortable loudness levels in pediatric patients with cochlear implants,'' Arch. Otolaryngol. Head Neck Surg. **127**, 956–960.
Hay-McCutcheon, M. J., Brown, C. J., Schmidt Clay, K., and Seyle, K. (**2002**). ''A comparison of EAP and EABR thresholds in Nucleus CI24R cochlear implant recipients,'' 25th Annual Midwinter Research Meeting of the Association for Research in Otolaryngology, St. Petersburg Beach, FL.
Lenarz, T., Tasche, C., Frohne-Buechner, C., Buechner, A., Illg, A., Strauss-Schier, A., and Battmer, R.-D. (**2001**). ''Comparison of different Clarion and Nucleus electrode systems within matched-pairs groups,'' 2001 Conference on Implantable Auditory Prostheses, Pacific Grove, CA.
Lesinski-Schiedat, A., Goldring, J. E., Gupta, S., Battmer, R.-D., Frohme, C., Weber, B. P., and Lenarz, T. (**2000**). ''Evaluation of Clarion Electrode Positioning System with the preformed and Hifocus electrode,'' CI2000: The 6th International Cochlear Implant Conference, Miami Beach, FL.
Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Miller, A. L., Morris, D. J., and Pfingst, B. E. (**1997**). ''Interactions between pulse separation and pulse polarity order in cochlear implants,'' Hear. Res. **109**, 21–33.
Osberger, M. J., Koch, D., Fisher, L., and Zimmerman-Phillips, S. (**1999**). ''Clinical results in patients implanted with the Clarion electrode positioning system,'' 1999 Conference on Implantable Auditory Prostheses, Pacific Grove, CA.
Parkinson, A. J., Arcaroli, J., Staller, S. J., Arndt, P. L., Cosgriff, A., and Ebinger, K. (**2002**). ''The Nucleus 24 Contour cochlear implant system: Adult clinical trial results,'' Ear Hear. **23**, 41S–48S.
Pfingst, B. E., Spelman, F. A., and Sutton, D. (**1980**). ''Operating ranges for cochlear implants,'' Ann. Otol. Rhinol. Laryngol. Suppl. **66**(89), 1–4.
Pfingst, B. E., and Morris, D. J. (**1993**). ''Stimulus features affecting psychophysical detection thresholds for electrical stimulation of the cochlea. II. Frequency and interpulse interval,'' J. Acoust. Soc. Am. **94**, 1287–1294.
Pfingst, B. E., Morris, D. J., and Miller, A. L. (**1995**). ''Effects of electrode configuration on threshold functions for electrical stimulation of the cochlea,'' Hear. Res. **85**, 76–84.
Saunders, E., Cohen, L., Aschendorff, A., Shapiro, W., Knight, M., Stecker, M., Richter, B., Waltzman, S., Tykocinski, M., Roland, T., Laszig, R., and Cowan, R. (**2002**). ''Threshold, comfortable level and impedance changes as a function of electrode-modiolar distance,'' Ear Hear. **23**, 28S–40S.
Shannon, R. V. (**1985**). ''Threshold and loudness functions for pulsatile stimulation of cochlear implants,'' Hear. Res. **18**, 135–143.
Shannon, R. V. (**1989**). ''A model of threshold for pulsatile electrical stimulation of cochlear implants,'' Hear. Res. **40**, 197–204.
Simmons, F. B. (**1966**). ''Electrical stimulation of the auditory nerve in man,'' Arch. Otolaryngol. **84**, 24–76.
Skinner, M. W., Holden, L. K., Holden, T. A., and Demorest, M. E. (**2000**). ''Effect of stimulation rate on cochlear implant recipients' thresholds and maximum acceptable loudness levels,'' J. Am. Acad. Audiol. **11**, 203–213.
Vandali, A. E., Whitford, L. A., Plant, K. L., and Clark, G. M. (**2000**). ''Speech perception as a function of electrical stimulation rate: Using the

Nucleus 24 Cochlear implant system,'' Ear Hear. **21**, 608–624.
Xu, L., and Pfingst, B. E. (**2002**). ''Variation in thresholds and comfort levels across cochlear implant stimulation sites: Effects of electrode configuration and stimulus level,'' 25th Midwinter Research Meeting, Association for Research in Otolaryngology, St. Petersburg Beach, FL.
Young, N. M., and Grohne, K. M. (**2001**). ''Comparison of pediatric Clarion recipients with and without the electrode positioner,'' Otol. Neurotol. **22**, 195–199.

Comment on "An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity" [J. Acoust. Soc. Am. 114, 2112–2117] (L)

Hendrikus Duifhuis[a)]
University of Groningen, BCN-Neuro Imaging Center and Department of Biomedical Engineering, Faculty of Mathematics and Natural Sciences, Antonius Deusinglaan 2, 9713AW Groningen, Netherlands

(Received 25 January 2004; revised 4 February 2004; accepted 5 February 2004)

This letter concerns the paper ''An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity'' [E. A. Lopez-Poveda, J. Acoust. Soc. Am. **114**, 2112–2117 (2003)]. It proposes a correction of the historical framework in which the paper is presented. Additionally it mentions some general problems associated with filter-bank models, and it discusses the validity and usefulness of the ''transfer function'' concept for nonlinear filters. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1694999]

PACS numbers: 43.66.Ba [WPS] Pages: 1889–1890

In a recent paper Lopez-Poveda (2003) appears to suggest that the dual-resonance nonlinear filter model goes back to Meddis *et al.* (2001). This suggestion I would classify as rather arbitrary and incomplete. The basic idea was introduced much longer ago.

Before giving some specific additional references it appears in order to emphasize the distinction between two classes of models of the cochlea. The first class consists of what is often called ''transmission line'' models, whereas the second class models in terms of filter banks. The major difference is that models in class 1 take physical coupling between system elements into account, whereas in class 2 the channels are independent, and coupling is completely determined by the common input. Obviously, in terms of system analysis the two types of model can be made to work identically as long as they are completely linear. However, as soon as a nonlinearity arises somewhere within the system, the two classes behave differently. It should be noted that if the ''elements'' of a transmission line model are second order, this does not imply that the equivalent transmission line elements are of the same order. Physical coupling of neighboring elements in the transmission line increases their order.

Both class 1 and class 2 contain models of different complexity. Starting from the simplest second order elements, the models are expanded by either feedback loops of various complexity or by parallel (feed forward) loops, usually in attempts to model parts of the organ of Corti structure and/or the tectorial membrane. Realistic modeling of the fine structure is far from complete, primarily because it is still extremely difficult to get reliable estimates of the physical parameters. Both classes have their physical, engineering, and mathematical elaborations.

Class 1 models go back to Zwislocki (1950) and Ranke (1950). Relevant contributions to the development of these models have been given by, e.g., Kim *et al.* (1980), de Boer (1980, 1984, 1991), Neely (1985), and Duifhuis (1988).

Class 2 models went through an almost simultaneous development. One of the earlier extensions is that of the second filter, or BPNL, introduced by the Washington University group around 1970 (e.g., Goblick and Pfeiffer, 1969; Pfeiffer, 1970). The proposal by Lopez-Poveda continues along that line, and is very similar to the network proposed by Goldstein (1988), and further developed in a number of subsequent papers. The objective of most class 2 models was to describe the cochlear transmission characteristics rather than the cochlear mechanics (e.g., Duifhuis, 1973, 1976; Zwicker, 1986). The more recent gamma-tone filter bank models also belong to class 2, and provide a time domain description which is very similar to the frequency domain filter bank proposed by Goldstein *et al.* (1971) and used by Duifhuis (1973). As referenced by Bleeck and Patterson in the current aim-manual (http://www.mrc-cbu.cam.ac.uk/cnbh/aimmanual): ''This function (i.e., the gamma-tone) was introduced by Aertsen and Johannesma (1980) and used by de Boer and de Jongh (1978) and Carney and Yin (1988) to characterize ''revcor'' data from cats.''

Finally, it is noted that it is dangerous to use the impedance concept for nonlinear systems. The ratio of output and input spectra cannot take the distortion products into account. Although a spectral analysis of (e.g.) overtone responses is of some use, it does not provide a general alternative for the impedance. In other words, for nonlinear systems there is no proper definition of impedance.

a)Electronic mail: duifhuis@bcn.rug.nl

Aertsen, A. M. H. J., and Johannesma, P. I. M. (**1980**). ''Spectro-temporal receptive fields of auditory neurons in the grassfrog,'' Biol. Cybernetics **38**, 223–234.
Carney, L. H. and Yin, T. C. T. (**1988**). ''Temporal Coding of resonances by low-frequency auditory nerve fibers: single-fiber responses in a population model,'' J. Neurophysiol. **60**, 1653–1677
de Boer, E. and de Jongh, H. R. (**1978**). ''On cochlear encoding: potentialities and limiations of the reverse-correlation technique,'' J. Acoust. Soc. Am. **63**, 115–135.
de Boer, E. (**1980**). ''Auditory Physics. Physical principles in hearing theory III,'' Phys. Rep. **62**, 87–174.
de Boer, E. (**1984**). ''Auditory Physics. Physical principles in hearing theory. II,'' Phys. Rep. **105**, 141–226.
de Boer, E. (**1991**). ''Auditory Physics. Physical principles in hearing theory. III,'' Phys. Rep. **203**, 125–231.

Duifhuis, H. (**1973**). ''Consequences of peripheral frequency selectivity for nonsimultaneous masking,'' J. Acoust. Soc. Am. **54**, 1471–1488.
Duifhuis, H. (**1976**). ''Cochlear nonlinearity and second filter: Possible mechanism and implications,'' J. Acoust. Soc. Am. **59**, 408–423.
Duifhuis, H. (**1988**). ''Cochlear Macromechanics,'' in *Auditory Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 189–211.
Goblick, Jr., T. J., and Pfeiffer, R. R. (**1969**). ''Time-domain measurements of cochlear nonlinearities using combination click stimuli,'' J. Acoust. Soc. Am. **46**, 924–938.
Goldstein, J. L. (**1988**). ''Updating Cochlear Driven Models of Auditory Perception: A New Model for Nonlinear Auditory Frequency Analysing Filters,'' in *Working Models of Human Perception*, edited by B. A. G. Elsendoorn and H. Bouma (Academic, London), pp 19–58.
Goldstein, J. L., Bear, T., and Kiang, N. Y. S. (**1971**). ''A theoretical treatment of latency, group-delay and tuning characteristics for auditory-nerve responses to clicks and tones,'' in *Physiology of the Auditory System*, edited by M. B. Sachs (National Educational Consultants, Baltimore), pp. 133–141.
Kim, D. O., Molnar, C. E., and Matthews, J. W. (**1980**). ''Cochlear mechanics: Nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure,'' J. Acoust. Soc. Am. **67**, 1704–1721.
Lopez-Poveda, E. A. (**2003**). ''An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity,'' J. Acoust. Soc. Am. **114**, 2112–2117.
Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (**2001**). ''A computational algorithm for computing nonlinear auditory frequency selectivity,'' J. Acoust. Soc. Am. **109**, 2852–2861.
Neely, S. T. (**1985**). ''Mathematical modeling of cochlear mechanics,'' J. Acoust. Soc. Am. **78**, 345–352.
Pfeiffer, R. R. (**1970**). ''A model for two-tone inhibition of single cochlear-nerve fibers,'' J. Acoust. Soc. Am. **48**, 1373–1378.
Ranke, O. F. (**1950**). ''Theory of operation of the cochlea: A contribution to the hydrodynamics of the cochlea,'' J. Acoust. Soc. Am. **22**, 772–777.
Zwicker, E. (**1986**). ''A hardware cochlear nonlinear preprocessing model with active feedback,'' J. Acoust. Soc. Am. **80**, 146–153.
Zwislocki, J. (**1950**). ''Theory of the acoustical action of the cochlea,'' J. Acoust. Soc. Am. **22**, 778–784.

Reply to "Comment on 'An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity' " (L)

Enrique A. Lopez-Poveda[a)]
Instituto de Neurociencias de Castilla y León, Universidad de Salamanca, Avenida de Alfonso X "El Sabio" s/n, 37007 Salamanca, Spain

(Received 9 February 2004; accepted for publication 10 February 2004)

[DOI: 10.1121/1.1695000]

PACS numbers: 43.64.Bt [WPS]

Duifhuis is correct to say that the paper of Lopez-Poveda (2003) lacks an adequate description of the historical framework that led Meddis *et al.* (2001) to develop and publish their dual-resonance nonlinear (DRNL) filter. Such a description was omitted in Lopez-Poveda (2003) for the sake of conciseness and because the paper was presented as an extension of the work of Meddis *et al.* (2001), where an historical account can be found. The historical review provided by Duifhuis in his comment letter is even more comprehensive and, therefore, most welcome.

Also welcome is Duifhuis' clarification concerning the concept of *transfer function* (hence of *impedance*). This, of course, does not apply to nonlinear systems like the DRNL filter and this is already acknowledged by Lopez-Poveda (2003). It is the impossibility of deriving an exact analytic transfer function for the DRNL filter that makes the *approximate* transfer function suggested by Lopez-Poveda (2003) such a useful tool for quickly evaluating the *peak*-amplitude and phase aspects of its response to pure tones. Furthermore, the approach of Lopez-Poveda (2003) may be generalized to derive approximate transfer functions for other nonlinear filter systems.

[a)]Electronic mail: ealopezpoveda@usal.es

Lopez-Poveda, E. A. (**2003**). "An approximate transfer function for the dual-resonance nonlinear filter model of auditory frequency selectivity," J. Acoust. Soc. Am. **114**, 2112–2117.

Meddis, R., O'Mard, L. P. O., and Lopez-Poveda, E. A. (**2001**). "A computational algorithm for computing nonlinear auditory frequency selectivity," J. Acoust. Soc. Am. **109**, 2852–2861.

Animations for visualizing and teaching acoustic impulse scattering from spheres

C. Feuillade[a)]
Naval Research Laboratory, Stennis Space Center, Mississippi 39529-5004

(Received 21 June 2002; revised 11 December 2003; accepted 30 December 2003)

The educational value of time–domain animations for visualizing acoustic impulse scattering from spheres, and the formation of scattered wave fronts, is demonstrated. Anderson's fluid sphere theory [J. Acoust. Soc. Am. **22**, 426–431 (1950)] is used to demonstrate scattering for two cases: (a) a fixed rigid sphere; (b) a pressure release sphere. The backscattering regime is seen to be dominated by geometric reflections. In the forward scatter region, visualizations for both cases show that the incident and forward scattered fields combine to rapidly minimize amplitude and phase perturbations of the wave front, leading to ''wave front healing.'' Diffraction into the acoustic shadow behind the sphere is seen and, in the rigid case, leads to a clearly discernible circumferential wave that breaks off in the backward direction. Animations based upon the raypath scatter method depict how this technique represents geometrical reflections, while omitting diffraction effects, but appears to be a reasonable approximation for backscattering applications. The Hickling and Wang ''movable'' rigid sphere theory [J. Acoust. Soc. Am. **39**, 276–279 (1966)] is used to create an animation that illustrates the ''rebound'' response motion of a sphere of finite mass to the action of the incident field. The animations discussed are downloadable via the World Wide Web. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1651116]

PACS numbers: 43.10.Sv, 43.20.Px, 43.20.Fn [VWS] Pages: 1893–1904

I. INTRODUCTION

In acoustics, as in other physical sciences, many processes and mechanisms are complex and subtle. From a didactic point of view, it is important to convey a solid understanding of the behavior exhibited when acoustic waves interact with objects or structures, causing them to move and vibrate. However, it is frequently difficult to understand the phenomena that occur, because they may not be readily visible, or may be hard to conceptualize. In teaching the physics of moving phenomena, in particular, any tool that can be used to visually, but accurately, demonstrate the complete ''cause and effect'' sequence of events that takes place during a complex interaction is clearly of significant educational value.

Scientific visualization using animation techniques has long been recognized as an extremely powerful method for conveying information, since it allows for the quick interpretation and comprehension of complex phenomena. Pioneering work in this area was performed by Bragg and his co-workers,[1] who developed a ''bubble-raft'' model for representing a crystal structure, and later produced a movie (i.e., a motion picture) based upon the model that graphically depicts the dynamical processes and mechanisms of dislocations. In acoustics, Neubauer,[2] and later Neubauer and Dragonette,[3] produced a series of papers that described the use of schlieren photography to produce still pictures and movies of wave interactions with cylinders. Hickling[4] used computers to produce contour plots for visualizing acoustic scattering from elastic spheres.

The use of computer animations in acoustics education is relatively new. In 1998, Sparrow and Russell[5] reported using *MATHEMATICA* to create short animation sequences of simple acoustical topics for use in their teaching work. Russell[6] later demonstrated how it was possible to use the mathematical power available with *MATHEMATICA* to generate animations to accurately represent complex wave phenomena in a visually assimilable form. Russell[7] has since developed a website containing many useful animation demonstrations for acoustics education.

In the past, the production of multimedia animations would have been an expensive, difficult, and time-inefficient task for most college and university educators, requiring specialized equipment and expertise. However, with the continued advance of computers, the World Wide Web, and the availability of sophisticated software at a reasonable, or even nominal, cost, it is becoming increasingly possible for most teachers to use relatively ''low-tech'' methods, which can be implemented in their office using a PC, to develop effective visualization tools. Sparrow and Russell primarily used *MATHEMATICA* for their computational work, afterward converting the animations into *QuickTime* movies. The present paper reports animations produced using *MATLAB* for the computational work, and then implementing *GraphicConverter* to concatenate the output graphics files into *QuickTime* movies, which can be played to demonstrate the temporal behavior. These procedures comprise just two examples of how multimedia demonstrations may be produced. As computer technology advances, the range of options available to the teacher for producing animations and other visualizations is undoubtedly proliferating.

The acoustic scattering properties of spheres, which comprise an important and useful special case in object scat-

[a)]Electronic mail: cf@nrlssc.navy.mil; telephone: 228-688-5783; fax: 228-688-5341.

tering, has historically received a lot of attention (Ref. 8, and see the list of references). However, while an acoustic scattering event is essentially a time–domain phenomenon, studies of scattering from objects have often concentrated on time-independent/cw (continuous-wave) descriptions, which often result in only a limited or partial understanding of the problem. This present work demonstrates how a comprehensive grasp of the scattering phenomenon may be facilitated by approaching the problem in the time–domain, and developing appropriate visualization techniques. The intention here is to provide accurate depictions of the predominant features of time–domain acoustic scattering from spheres, which can be used to give students and others a clear understanding of the physical processes involved, and to guide future thinking about other related scattering problems.

In this work, time–domain visualization animations of impulse scattering from rigid and pressure release spheres are presented. The computations are based upon (a) Anderson's fluid sphere scattering theory;[9] (b) the raypath scatter impulse wave solution (Ref. 10, Sec. 7.5.1); and (c), the Hickling and Wang ''movable'' rigid sphere scattering theory.[11] Damping processes are not incorporated in the animations shown here, since the purpose is to demonstrate the time domain evolution of the scattered field for the classical problem of impenetrable spheres. In the case of Anderson's theory, calculations over a range of physical locations and temporal frequencies are made, and a Fourier synthesis technique is used to produce representations that may be animated in the time–domain. Anderson's sphere scattering theory employs a ''full wave'' representation, and provides physically complete and accurate depictions of the scattering phenomena. The computational results lead to movie animations that enable important insights to be gained into the physical processes that lead to the construction of scattered wave fronts. It is possible to time separate and identify different scattering mechanisms, including geometrical scatter; diffraction; and wave interference. A phenomenon of significant interest is the interaction of wave fronts in the forward scatter direction. The animations clearly depict how, in this region, the incident field and the forward scattered field interfere to ''fill in'' the acoustic shadow cast by the sphere, leading to the phenomenon called ''wave front healing.''[12,13] This term expresses the fact that travel time and amplitude anomalies in the forward scattering region do not keep their initial values (attained just as the wave front passes the object), but diminish as the wave progresses in space.

The animations based upon the raypath scatter method depict how this impulse solution accurately represents the geometrical scatter of a sphere, but omits the effects of diffractions and the associated wave interactions. It is educationally very helpful to compare and contrast these animations with those produced using Anderson's theory for the equivalent cases. It becomes clear that, for many applications that require just the backscattered wave component, the raypath scatter method provides a very reasonable approximation for solving the scattering problem. The animation produced using the Hickling and Wang theory is presented to illustrate the phenomenon of the ''rebound'' response motion of a sphere of finite mass to the action of the incident field.

The primary purpose of this paper is to introduce, and provide a commentary on and analysis of the movie animations presented for the various cases of acoustic scattering from spheres. The figures shown in the paper depict single frames, or ''snapshots,'' excerpted from the movies, and are intended only to introduce, not substitute for, the content of the movies. To receive the full didactic benefit of the animations, they should be viewed for themselves. The reader is encouraged to go to the website indicated in Ref. 14, where the movies are freely downloadable. Educators are welcome to use the movies for their own teaching purposes, and the author welcomes any feedback they may have.

The next section presents and describes the animations made using Anderson's theory, the raypath scatter method, and the Hickling and Wang theory. This is followed by a brief summary of conclusions.

II. THE ANIMATIONS

A. How the animations were made

All of the scientific computations for the work described in this paper were performed using *MATLAB*.[15] As an intrinsic part of the computer algorithms, appropriate graphical displays were chosen to display the results of the calculations. For example, the acoustic pressure amplitude for a grid of locations surrounding a spherical scatterer at a particular point in time was calculated, and then the pressure values were plotted in the form of a two-dimensional (2-D) color shade plot, with a corresponding color bar included for reference. Numerous individual plots were created for a sequence of equally spaced time points spanning the scattering event. Each of these plots was individually saved in pict format (in some cases they were saved in eps format, and later converted to a pict format using external software). When all of the shade plots for a specific physical case had been made, they were then processed into a *QuickTime* animation using the commercially available software package *GraphicConverter*, which is able to concatenate the individual frames together in the correct sequence to form a single unbroken movie.

Five of the animations presented in this work involve color shade plots. To make the graphics, considerable attention was given to the choice of color scale. A recommendation was followed, made by Sparrow *et al.*,[16] of using a ''red–blue'' color scale. The simple partition into red (positive field amplitude) and blue (negative field amplitude) allows regions of positive and negative pressure features to be readily distinguished in the movies, and also enables color graphics to be made that are intelligible to the largest number of individuals, including those with many common forms of color deficiency.

B. Animations produced using Anderson's theory

Anderson's scattering theory, full details of which may be found in the literature [Refs. 9 and 17 and Ref. 10, pp. 276–280], describes acoustic scattering by a fluid-filled sphere entrained within a second (different) fluid. The acoustic field inside the fluid sphere consists of compressional waves, as do the incident and scattered fields in the external

fluid. The density and sound speed in the internal and external fluids can be independently varied. In water (to take just one field of application) this feature has enabled Anderson's theory to be used successfully to describe scattering from objects as different as air bubbles (high contrast) (Ref. 10, pp. 289–291), and zooplankton (low contrast),[18] over a broad frequency range. In this present work, Anderson's theory is used to develop educational demonstrations of the time domain evolution of the scattered field for the classical problems of (a) a fixed rigid (impenetrable) sphere; and (b) a pressure release (impenetrable) sphere.

Anderson's method utilizes the axial and spherical symmetry of the scattering problem by expanding the internal and external acoustic fields in terms of spherical modes expressed as products of spherical Bessel/Hankel functions and Legendre polynomials. The geometric origin is placed at the center of the sphere. Let the external fluid have density ρ, sound speed c, and propagation wave number $k=2\pi f/c=2\pi/\lambda$, where f is the frequency of the incident field, and λ is the wavelength in the external fluid. The corresponding parameters for the internal fluid are ρ_1, c_1, and $k_1=2\pi f/c_1=2\pi/\lambda_1$. The ratios of these parameters are written as $g=\rho_1/\rho$, $h=c_1/c$, and $k_1=k/h$. Let the radius of the sphere be R. An incident plane wave of amplitude $\mathbf{P}$ traveling in the $+z$ direction (such that backscatter would be observed at $\theta=\pi$) is expanded as

$$P_{\text{inc}}(f)=\mathbf{P}\sum_{m=0}^{\infty} i^m(2m+1)P_m(\cos\theta)\, j_m(kr)\,, \tag{1}$$

where the field is evaluated at (r,θ). The scattered field ($r>R$) is written as

$$P_s(f)=\sum_{m=0}^{\infty} A_m\, P_m(\cos\theta)\, h_m(kr)\,, \tag{2}$$

and the field inside the sphere ($r\leqslant R$) is

$$P_{\text{int}}(f)=\sum_{m=0}^{\infty} B_m\, P_m(\cos\theta)\, j_m(k_1 r). \tag{3}$$

In these expressions, A_m, B_m are the coefficients of the mth partial wave for the scattered and internal fields, respectively, $P_m(\cos\theta)$ is the mth Legendre polynomial, and $j_m(kr)/h_m(kr)$ the corresponding spherical Bessel/Hankel functions. To determine A_m, two boundary conditions are applied at the surface of the sphere: the continuity of the normal displacement and the continuity of normal pressure, yielding[9]

$$A_m = -\mathbf{P}\, i^m\,(2m+1) \times\frac{j_m'(k_1R)\, j_m(kR)-gh\, j_m(k_1R)\, j_m'(kR)}{j_m'(k_1R)\, h_m(kR)-gh\, j_m(k_1R)\, h_m'(kR)}, \tag{4}$$

which may be directly substituted in (2) to determine $P_s(f)$.

The scattered impulse response is formally given by the inverse Fourier transform

$$p_s(t)=\int_{-\infty}^{\infty} P_s(f)\, e^{i2\pi ft}\, df\,. \tag{5}$$

The impulse responses for the spheres in this analysis were computed by evaluating (5) numerically using the inverse discrete Fourier transform (via the inverse FFT algorithm), i.e.,

$$p_s(n\,\Delta t)=\frac{1}{M}\sum_{m=0}^{M-1} P_s(m\,\Delta f)\; W_s(m\,\Delta f)\; e^{i2\pi mn/M}, \tag{6}$$

where M is the number of points in the inverse FFT. A similar procedure was used by Hickling[19] to compute the pulse forms of echoes backscattered from elastic spheres in water. In this work, the values of $P_s(m\,\Delta f)$ were computed by means of Anderson's theory for a range of receiver locations around the sphere, using values of A_m determined for the specific cases under consideration. An algorithm was written to sort the received pressure at these locations according to time index, and thus produce a causal sequence of 2-D color plots. These were then concatenated together using *Graphic-Converter* to produce *QuickTime* movies. The factors $W_s(m\,\Delta f)$, appearing on the right-hand side of (6), represent a shading window applied to the spectral amplitudes in the frequency domain, which has the effect of reducing Gibbs oscillations in the time signal produced by the inverse FFT. A Blackman shading window was used here. The temporal "width" of the impulse signals in the animations can be reduced by increasing the frequency bandwidth in the inverse FFT. However, this also increases the computational cost. In the cases studied here, the bandwidth ranges from $kR[=(2\pi f/c)R]=0$ to $kR=34.56$, with a sampling interval $\Delta(kR)=(2\pi\,\Delta f/c)R=0.135$. At the high-frequency end, $R/\lambda\approx 5.5$, which provides sufficient resolution to accurately represent the wave phenomena in the animations. Further calculations (not shown), performed using larger FFT bandwidths, indicated no significant differences in the display of the wave processes for the cases discussed in this work.

1. Example: Fixed rigid sphere

In Anderson's paper, scattering from a fixed rigid sphere is represented by allowing the ratios g and h to approach infinity. The coefficients A_m appearing in the expansion (2) for the scattered field then have the following asymptotic form:

$$\lim_{g,h\to\infty} A_m=-i^m\,(2m+1)\,\frac{j_m'(kR)}{h_m'(kR)}\,, \tag{7}$$

where it is assumed that $\mathbf{P}=1$. This "infinite density" limitation of Anderson's description of scattering for a rigid sphere was removed in 1965 by Hickling and Wang,[11] who determined the scattering response of a movable rigid sphere, and so allowed the behavior of rigid objects of finite density to be described. Anderson's prescription is used here, however, to study scattering and wave front interactions independently of the translational rebound response of the scatterer. The rebound response is examined in a later section.

In Fig. 1(a), the solid black line shows the farfield backscattered pressure impulse response for a fixed rigid sphere, calculated from Anderson's theory, using the coefficients prescribed by (7) in the field expansion (2). The signal amplitudes of the scattered field are referred to a range of one

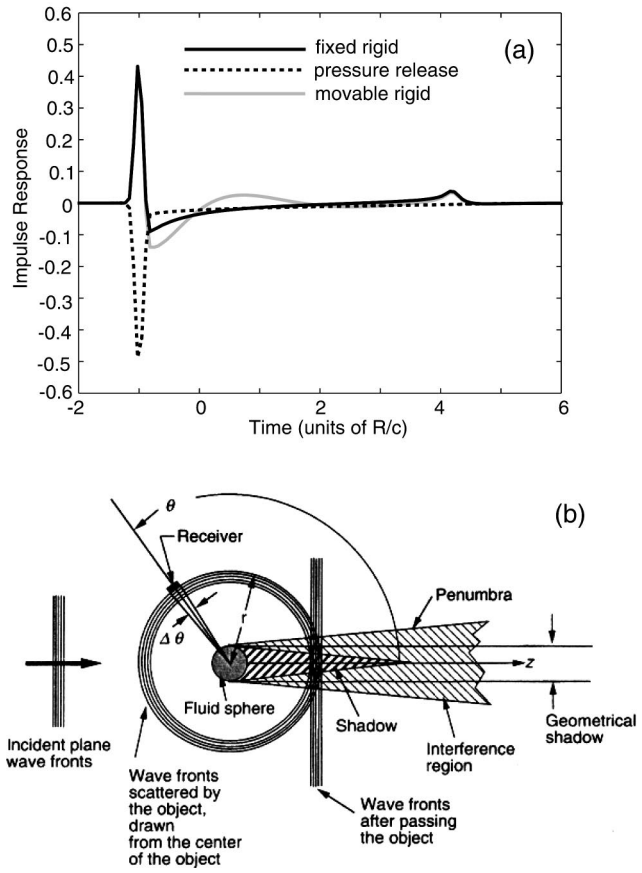


FIG. 1. (a) Backscattered pressure impulse responses of a fixed rigid sphere (solid black line), a pressure release sphere (dashed line), and a movable rigid sphere (gray line). The time axis is scaled in units of R/c; (b) scattering geometry.

radius from the center of the sphere, and scaled to the peak $(=1)$ of the incident field impulse. The time axis is scaled in units of R/c, and referenced so that the incident plane wave pulse passes the center of the sphere at $t=0$. The first (leftmost) feature seen at $t=-R/c$ is a sharp positive spike, which corresponds to geometrical scatter from the front face of the sphere. For a fixed rigid sphere the geometric scattering length$=R/2$ (Ref. 10, pp. 268–271). When referred to one radius, this corresponds to $p_s(\text{geom})=0.5$. The geometric spike in Fig. 1(a) has a peak value ≈ 0.43, which is less than 0.5, due to the computational limitation of the bandwidth used in the inverse FFT. [For a further discussion of the magnitude of impulsive functions under these conditions, see Feuillade *et al.* (Ref. 20, Appendix B)]. The initial spike is followed by a smooth, slightly negative, response, which gradually rises to a small positive peak at slightly greater than $t=4R/c$. This feature is due (as will be clear from the accompanying animation in Fig. 2) to a circumferential wave propagating around the sphere.

Figure 2 shows selected frames from an animated movie depicting the scattering of an incident impulsive plane wave from a fixed rigid sphere, calculated using the same prescription used to produce the solid black line in Fig. 1(a). The pressure field is plotted, and scaled throughout to the peak value $(=1)$ of the incident field. The scatterer is represented by a black disk, since no field is generated inside the sphere, and the axes are scaled in radii. The sequence of frames enables the different physical mechanisms to be identified. The wave is incident from left to right. The first interaction seen [Fig. 2(a)] is geometric scattering from the front face. This corresponds to the initial positive spike in the solid black line in Fig. 1(a). Behind the backscattered wave front, the color scale shows that there is a region of negative pressure, as also seen in the solid black line of Fig. 1(a). As the incident wave passes by, it casts a partial acoustic shadow behind the sphere. This is ''filled in'' by diffraction from the incident wave front, and leads to a time-delayed portion of the wave front immediately behind the sphere [Fig. 2(b)]. As time proceeds further, diffraction continues to fill the shadow, the time delay disappears, and the wave front is ''healed''[12,13] [Fig. 2(c)], such that, as the advancing wave front propagates beyond the sphere in the forward scatter direction, the wave front shows progressively smaller effects of its encounter with the sphere. In addition, diffraction into the shadow behind the sphere is also seen to initiate a circumferential wave, which propagates around and breaks off in the back direction [Figs. 2(b)–2(d)]. This is the origin of the small positive peak (at $t\approx 4R/c$) in the solid black line of Fig. 1(a). Note carefully how the educational value of the movie animation is clearly exhibited. It illustrates the different physical processes, i.e., geometrical scatter, diffraction, and interference, coming into play during the interaction with the sphere, and enables the features seen in the solid black line of Fig. 1(a) to be given a meaningful physical interpretation. For example, the time difference between the geometric spike and the circumferential wave peak in Fig. 1(a), i.e., $\Delta t\approx 4R/c-(-R/c)=5R/c$, can now be interpreted as the time taken for the incident wave to travel from the front face of the sphere to the center plane of the sphere (R/c), plus the time taken for the field to diffract 180° around the rear face of the sphere $(\approx 3R/c)$, plus the time taken for the circumferential wave to propagate in the backscattering direction from the center plane of the sphere to the front face (R/c).

A precise numerical implementation of Anderson's fluid sphere scattering model was used recently[17] to investigate cw forward scatter phenomena. While the magnitude of the forward scattered field for spheres is often greater than the backscattered field, especially at high frequencies, this forward scattered component interferes destructively with the incident field to create a ''shadow'' region. Figure 1(b) depicts how experimental measurements of directional sound scatter are typically made. Simple time domain measurements could be used to separate the scattered sound from an incident ping when they do not arrive at the same time. When the receiver is moved to the forward scatter region, the incident and scattered fields arrive at nearly the same time, and interfere. In the case of plane wave incidence, the amplitude of the field does not change with range. However, the scattered field amplitude decreases with distance from the sphere. Since the total external pressure is the sum of the two, at very large ranges the pressure tends to the incident field (hence the term ''wave front healing''[12,13]). Close to the sphere, the two components interfere to create an acoustic shadow.

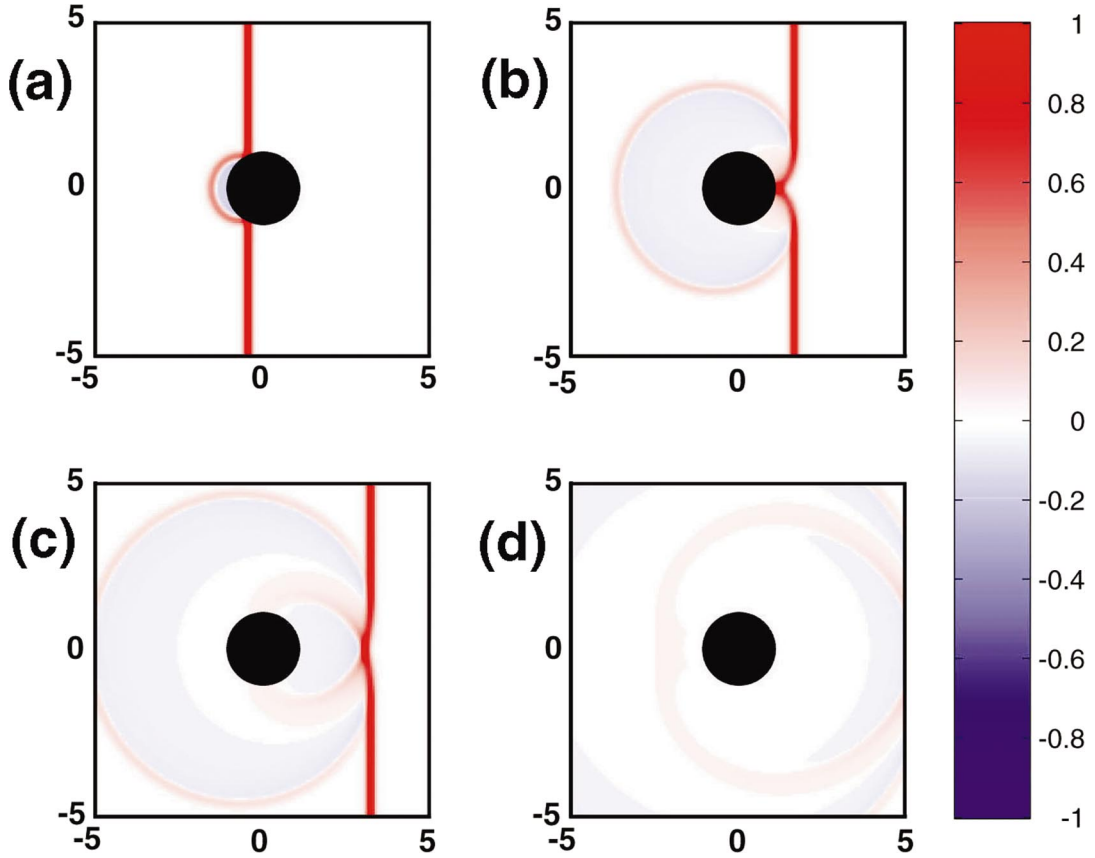


FIG. 2. Impulse scattering from a fixed rigid sphere. The pressure field is plotted, and scaled to the peak value (=1) of the incident field. The scatterer is represented by a black disk, since no field is generated inside the sphere. The axes are scaled in radii R. This figure was produced using Anderson's fluid sphere scattering theory.

Figure 3 shows frames from a movie animation that depicts, in a different form, the wave processes in the forward scatter direction. Three curves are shown. The solid blue line represents the incident acoustic pressure field (scaled to peak value=1). The dashed red line is the forward scattered pressure field. The solid black line is the total pressure field (i.e., the sum of the incident and scattered fields). The ordinate axis represents time (scaled in units of R/c), and the abscissa axis represents signal amplitude. Each frame of this figure [i.e., Figs. 3(a)–3(d)] depicts the time variations of the three signals that would be detected by a receiver placed at a specific location on the z axis, at a given distance from the sphere, in the forward scatter direction. As the movie plays, the receiver is progressively moved from its initial position directly adjacent to the rear face of the sphere [Fig. 3(a)], to locations farther from the sphere along the z axis [Figs. 3(b)–3(d)]. The receiver location, scaled in units of radii, is noted at the bottom of each frame. For display purposes, the curves are shifted back in time, so that the incident field peak is always depicted at time $t=0$.

The first frame of Fig. 3 shows the time variation of the field at $z=1R$. We see that at $t=0$, the scattered field exactly cancels the incident field to give zero total field. This represents the initial acoustic shadow immediately behind the sphere. At time about $t=0.6R/c$ later, however, the scattered (and, since the incident field is zero at this point, total field) registered at $z=1R$ rises to a peak, as sound is diffracted from other locations on the incident wave front to fill the acoustic shadow (cf. Fig. 2). As z increases, the time delay between the incident peak, and the peak in the total field, rapidly decreases. Consequently, even though the total wave field on the z axis is noticeably delayed behind the incident wave field at locations close to the sphere, the time delay shrinks very quickly as the wave continues to propagate forward (cf. Fig. 2 again). We also see that the amplitude of the incident field does not change as the range z increases. The scattered field amplitude, however, decreases with range. Since the total field results from adding the scattered field to the incident field, this implies that both the time delay and amplitude difference between the incident field and the total field quickly diminish. When $z=80R$, the incident and total fields are almost indistinguishable. The animation represented by the curves shown in Fig. 3 demonstrate the phenomenon of wave front healing. While a finite amount of energy is clearly removed from the incident wave field by the sphere, and scattered in other azimuthal directions, diffraction processes are observed to cause a redistribution of energy from other parts of the incident wave to fill the resulting shadow. Before the incident wave has propagated very far beyond the sphere, the energy loss is shared along a large portion of the remaining wave front, so that the amplitude difference at any one location is minimized. The phenomena that occur in the forward scattering region, and particularly the interaction between the forward scattered field and the incident field, has tended to be a subject of confusion in the past. The animations represented in Figs. 2 and 3 enable the

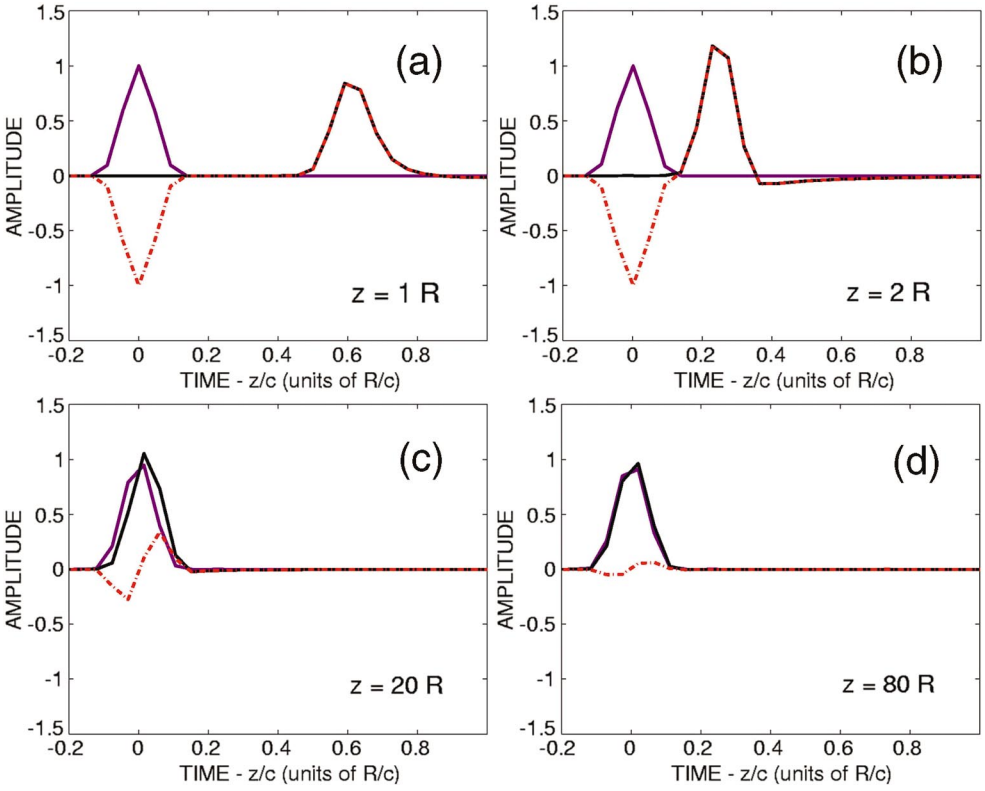


FIG. 3. Forward scattering processes of a fixed rigid sphere. The amplitudes of the incident pressure field (solid blue line), the scattered pressure field (dashed red line), and the total pressure field (solid black line), are plotted as a function of time. The amplitudes are scaled to the peak (=1) of the incident field. The four figures show the field variations detected at four different receiver placements. The incident field peak is always placed at $t=0$. The small variations in the shape of the peak tops are an artifact of the FFT sampling interval used to make the figure.

student to achieve a much improved understanding of the manner in which the fields combine to, first, create the acoustic shadow, and, second, minimize the effects of travel time and amplitude anomalies in the forward scattering region as the wave continues to propagate.

2. Example: Pressure release sphere

In Anderson's theory, the application of the pressure release boundary condition (which is almost physically realized, for example, in the case of an air bubble in water) yields the following form for the coefficients A_m in the expansion (2) for the scattered field:

$$A_m = -i^m(2m+1)\frac{j_m(kR)}{h_m(kR)}. \tag{8}$$

In Fig. 1(a), the dashed line is the farfield back scattered pressure impulse response for a pressure release sphere. Note that the geometric scattering spike is negative. This is a result of the boundary condition. Since the pressure at the boundary is forced to zero, a negative-going spike must be generated to cancel the positive incident field at the surface of the sphere. While the spike is negative, however, we see that $|p_s(\text{geom})| \approx 0.5$ (i.e., it still approaches the geometric scattering length$=R/2$). In contrast to the solid black line in Fig. 1(a) (the rigid sphere case), the dashed line contains no discernible feature corresponding to a circumferential wave. Uberall and his co-workers[21,22] have shown that, while a circumferential wave is expected in the pressure release case, its amplitude should be much less than for a rigid sphere. The limitations of the linearly scaled visual display used in Fig. 1(a) (and in the corresponding animations) preclude the appearance of extremely small features such as a possible circumferential wave in the pressure release sphere case.

Figure 4 shows frames from an animation depicting the scattering of an incident impulsive plane wave from a pressure release sphere, calculated using the coefficients given in Eq. (8). The pressure field is plotted, and scaled throughout to the peak value (=1) of the incident field. No field is generated inside the sphere, which, this time, is represented by a gray disk. The first interaction seen [Fig. 4(a)] is geometric scattering from the front face, but the shading indicates it is negative-going (see the color scale). This is extremely helpful for the student since it clearly shows, as we saw in Fig. 1(a), that a negative field is generated to cancel the positive incident field at the surface of the sphere, to fulfill the pressure release boundary condition. The geometric field propagates outward as time increases, and remains negative, but is reduced in amplitude. As the incident wave passes it casts a partial acoustic shadow behind the sphere, just as for the fixed rigid sphere, giving rise to a time-delayed section of the wave front [Fig. 4(b)]. This again is filled in by diffraction from the incident wave front, and eventually the time delay disappears as the wave front is healed [Fig. 4(c)]. In agreement with Fig. 1(a), diffraction into the shadow area behind the sphere does not give rise to a discernible circumferential wave in this animation.

Figure 5 shows movie frames depicting the evolution of the incident pressure field (solid blue line), scaled to peak

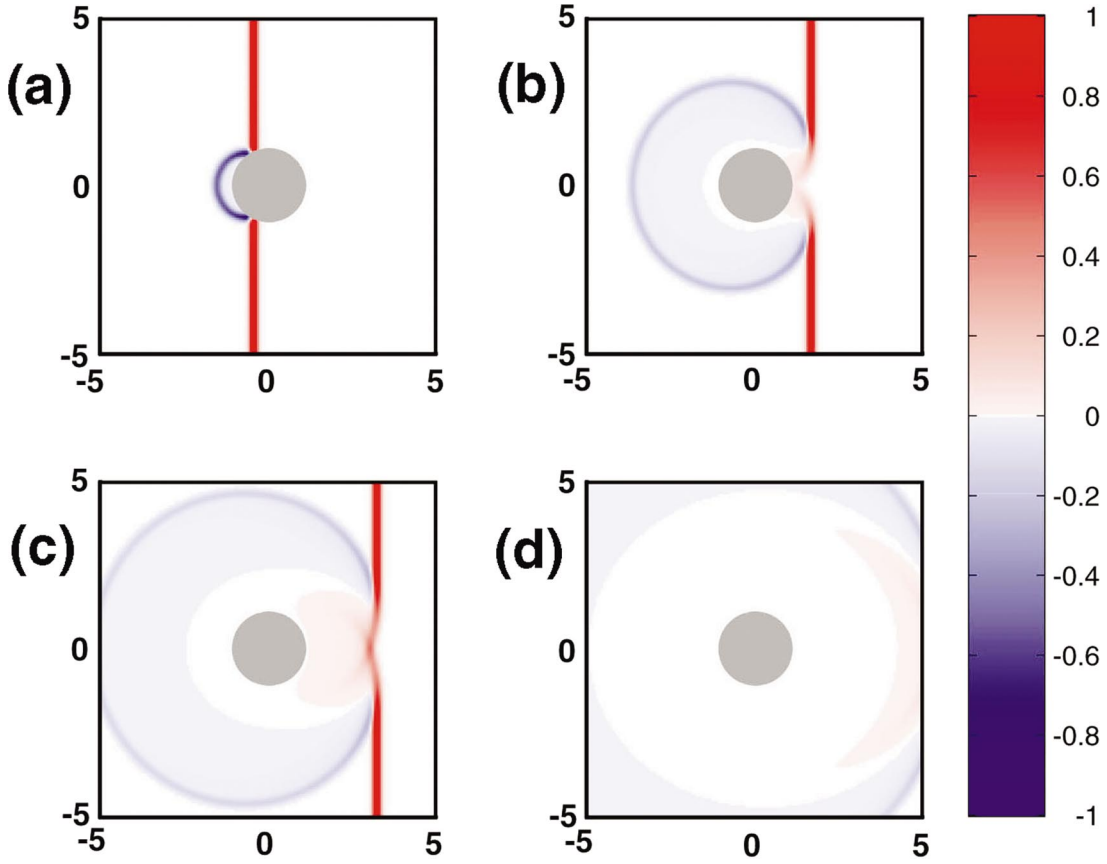


FIG. 4. Impulse scattering from a pressure release sphere. The pressure field is plotted and scaled to the peak value (=1) of the incident field. No field is generated inside the sphere, which in this case is represented by a gray disk. The axes are scaled in radii R. This figure was produced using Anderson's fluid sphere scattering theory.

value =1, the forward scattered pressure field (dashed red line), and the total pressure field (solid black line), for the pressure release case. The movie shows, as in Fig. 3, the effect of moving the receiver location from $z=1R$ to $z=80R$ along the z axis. The first frame shows the time variations at $z=1R$. The scattered field at $t=0$ again cancels the incident field to give a zero total field, but diffraction behind the sphere does not give rise to a time-delayed peak in this case. This is due to the pressure release boundary condition. In this case, we can clearly see that the field at $z=1R$ (i.e., precisely at the boundary of the sphere) is forced to zero *at all times*. However, as z moves away from the pressure release boundary, a nonzero delayed total field is generated, which quickly grows as z increases. As in the fixed rigid sphere case, diffraction causes the time delay and amplitude of the total field to rapidly approach that of the incident field as the wave front propagates forward. The animations represented by Figs. 4 and 5 show the student that the interaction of the scattered and incident fields in the forward scattering region for a pressure release sphere leads to the same "wave front healing" phenomenon as for a rigid sphere, even though the boundary conditions for the two objects are very different.

C. Animations produced using the raypath scatter method

Raypath calculations are one of the primary tools taught to students for describing acoustic propagation and scattering phenomena. In this section, movies are presented showing scattering from spheres as calculated using the raypath method. The purpose here is to contrast and compare the raypath method with the full wave method represented by the Anderson theory, for the rigid and pressure-release cases, with the intention of investigating the regime of validity of the raypath approximation.

The raypath expressions for the scatter of an incident plane wave by a rigid sphere have been in the literature for many years (see Ref. 10, Sec. 7.5.1, for a recent exposition). In order to develop movie animations showing the geometrical scatter of the acoustic pressure field from a rigid sphere, it is necessary to obtain explicit expressions for both the ray travel times, and the field amplitudes, for receiver locations around the sphere.

In the Kirchhoff approximation, reflections from an area on the boundary of the sphere are treated as if the local curved surface is a plane. Using a ray picture, as shown in Fig. 6, scattering from the sphere consists of a spray of reflected rays that follow the local Kirchhoff approximation. Each ray is reflected with its angle of reflection equal to its angle of incidence, as if the reflection point lies on a plane that is tangent to the sphere at that point. Diffraction effects are ignored in this approximation.

The scattering of impulsive wave fronts from a fixed, rigid, perfectly reflecting sphere may be calculated in the following manner. Let the incident sound be an impulsive

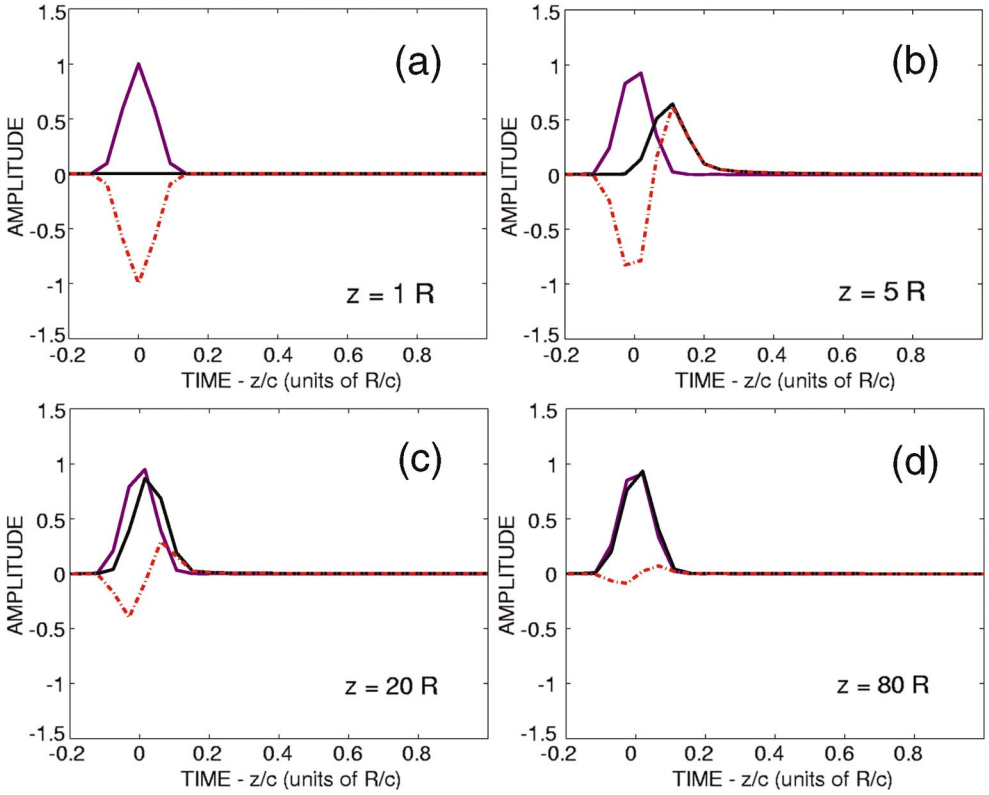


FIG. 5. Forward scattering processes of a pressure release sphere. The amplitudes of the incident pressure field (solid blue line), the scattered pressure field (dashed red line), and the total pressure field (solid black line), are plotted as a function of time. The amplitudes are scaled to the peak (=1) of the incident field. The four figures show the field variations detected at four different receiver placements. The incident field peak is always placed at $t=0$. The small variations in the shape of the peak tops are an artifact of the FFT sampling interval used to make the figure.

plane wave $p_0(t)$, and define the incoming energy density as

$$I_0=\int |p_0(t)|^2\,dt. \tag{9}$$

Assume that there is no energy absorption in the external medium, and no energy penetrates into the sphere. In this case, the scattered sound pressure at the receiver location (r,θ), at range r_{23} from the surface of the sphere (see Fig. 6), is proportional to a delayed incident sound pressure. Using Fig. 6, the time delay is referred to the incident plane wave contacting the top of the sphere, and the scattered sound is proportional to the incident impulse wave, i.e.,

$$p_{23}[t-(r_{23}+R(1-\cos\theta_1)/c]\propto p_0(t). \tag{10}$$

The incident energy on the surface area increment $dS=a^2\sin\theta_1\,d\theta_1\,d\phi$ of the sphere's surface is

$$E_0=\frac{a^2}{2}\sin 2\theta_1\,d\theta_1\,d\phi\int |p_0(t)|^2\,dt. \tag{11}$$

The rays falling within the angular increment $d\theta_1$, at angle θ_1, are scattered within the increment $d\theta_2=2d\theta_1$, at angle $\theta_2=2\theta_1$, and appear to originate from a "source" placed at $\tilde{S}$ (see Fig. 6). The scattered energy, measured at the point (r,θ), and range r_{23}, is

$$E_{23}=(r_{23}+r_{12})^2\sin\theta_2\,d\theta_2\,d\phi\int |p_{23}(t)|^2\,dt, \tag{12}$$

and assuming there is no loss of energy, $E_{23}=E_0$, so that Eq. (11) and Eq. (12) may be equated to yield

$$\int |p_{23}(t)|^2\,dt=\frac{a^2}{4(r_{23}+r_{12})^2}\int |p_0(t)|^2\,dt. \tag{13}$$

The proportionality in Eq. (10) then leads to

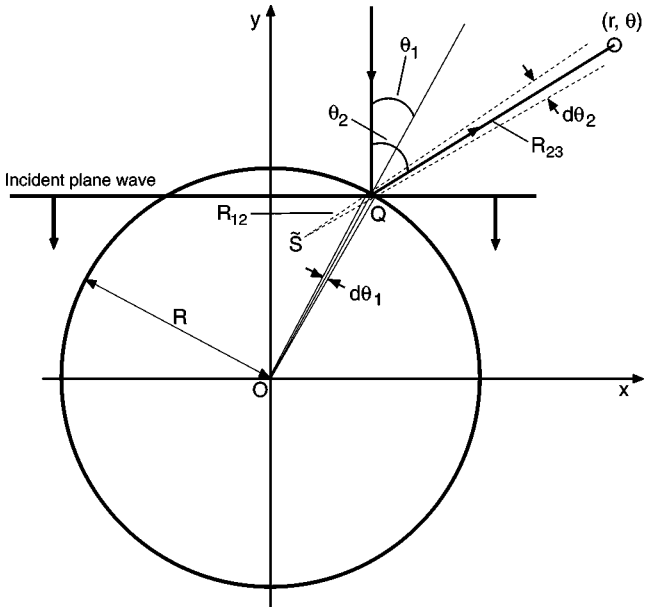


FIG. 6. Raypath acoustics construction for geometric scatter from a fixed rigid sphere, where $kR\gg 1$.

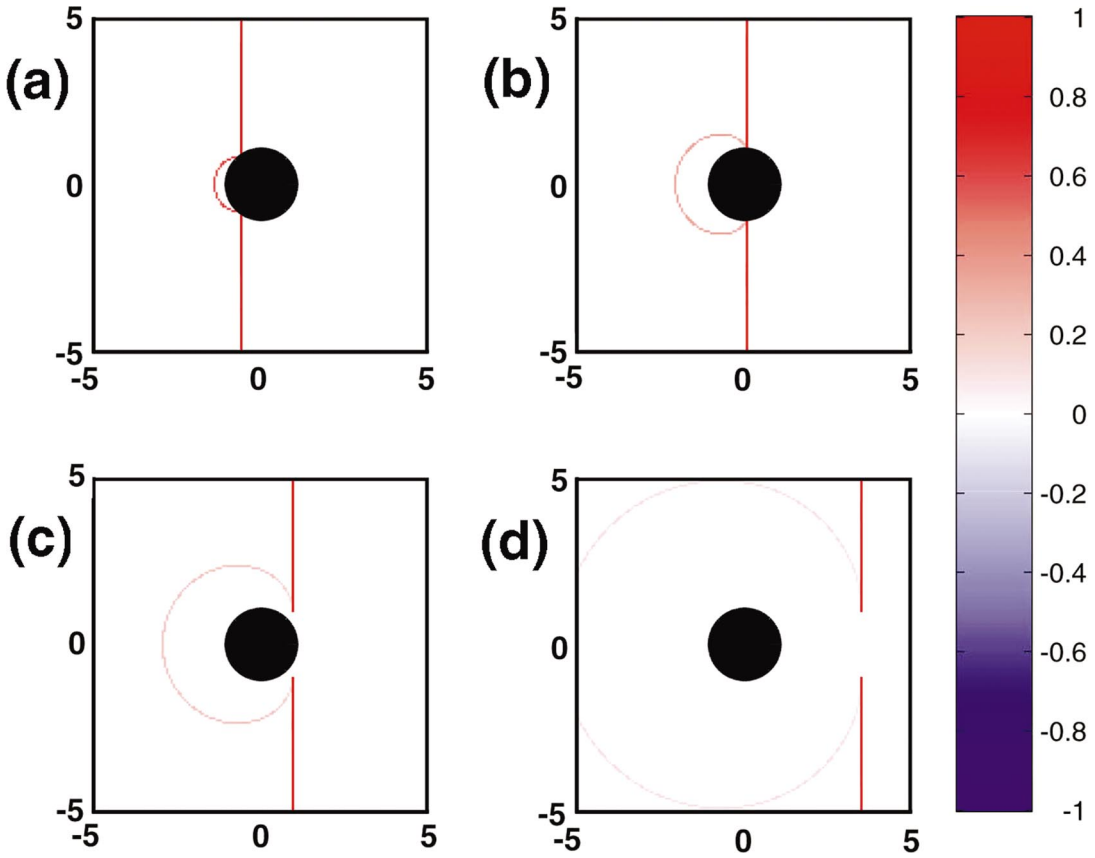


FIG. 7. Geometric impulse scattering from a fixed rigid sphere. The pressure field is plotted, and scaled to the peak value (=1) of the incident field. The scatterer is represented by a black disk, since no field is generated inside the sphere. The axes are scaled in radii R. This figure was produced using the raypath scatter method.

$$p_{23}[t-(r_{23}+R(1-\cos\theta_1)/c]=\frac{a}{2(r_{23}+r_{12})}p_0(t). \tag{14}$$

Equation (14) is a simplified ray trace solution for the scatter of sound by a rigid sphere. It does not give the diffracted components that are behind the sphere, nor the circumferential wave that travels around the sphere. The scattered sound pressure depends on the angle $\theta_2(=2\theta_1)$, because it depends on $(r_{12}+r_{23})$, the range from the apparent source $\tilde{S}$. The value of θ_1 corresponding to the receiver location (r,θ) may be determined by solving the equation

$$\sin(\theta-2\theta_1)=\frac{R}{r}\sin\theta_1. \tag{15}$$

Once θ_1 has been found, the range from $\tilde{S}$ to the boundary of the sphere is given by $r_{12}=(R\cos\theta_1)/2$. By computing values of p_0, p_{23}, and the time delay $[t-(r_{23}+R(1-\cos\theta_1)/c]$, for a range of receiver locations around the sphere, and values of t spanning the scattering event, causal sequences of 2-D plots for the total pressure were produced, and converted into *QuickTime* animations using *GraphicConverter*. Unlike the animations produced using the Fourier transform method in the Anderson theory cases, the wave front resolution achievable using the raypath scatter method is not bandwidth limited, but is determined by the grid spacing chosen for the 2-D plots.

Figure 7 shows selected frames from an animated movie depicting the scattering of an impulsive plane wave from a rigid sphere, calculated using the raypath scatter method. The maximum amplitude of the incident plane pressure wave $p_0=1$. The total field pressure p_0+p_{23} is plotted, and the sequence of frames enables the process of geometric scattering from the sphere to be followed. It is very helpful for a student to compare Fig. 7 with Fig. 2. Both the Anderson theory and the raypath scatter method clearly depict geometrical scatter from the surface of the sphere. Detailed numerical comparisons (not shown here) of data from the computations used to create the animations indicate that the expanding radial distributions predicted for the geometric wave front as a function of time are essentially identical for the two cases. This is the case for all azimuthal directions except those within the exact geometric shadow region (where the raypath scatter method does not calculate a wave front). Corresponding comparisons of the amplitude variations of the geometric wave front indicate that the raypath scatter calculations predict amplitudes that are ~10%–15% greater than those of the Anderson calculations for the same locations. The difference appears to be due to the limited FFT bandwidth used in the Anderson case. Within, or close to, the forward scattering region [which may be defined as the forward azimuthal cone delimited by the points at which the geometric scattering wave front rejoins the incident wave front (see Figs. 2 and 7)], the differences between the two methods of calculation are much greater.

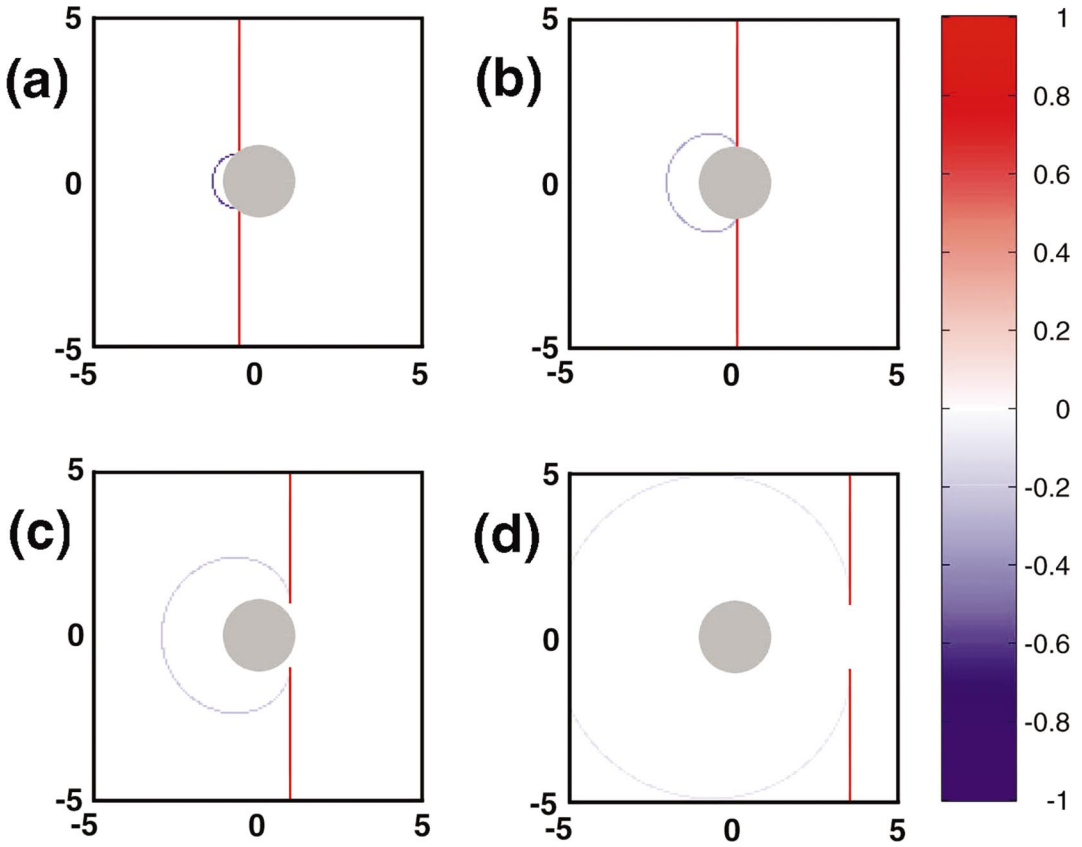


FIG. 8. Geometric impulse scattering from a pressure release sphere. The pressure field is plotted, and scaled to the peak value (=1) of the incident field. No field is generated inside the sphere, which in this case is represented by a gray disk. The axes are scaled in radii *R*. This figure was produced using the raypath scatter method.

Notably absent from Fig. 7 are the effects of diffraction. First, we see that the region of negative pressure directly following the geometrical wave front, seen in Fig. 2 [also in Fig. 1(a)], is not predicted by the raypath method. Second, the phenomenon of diffraction from the incident wave *behind* the sphere, to fill the acoustic shadow and heal the wave front, does not occur. As the wave front advances beyond the sphere, the geometric shadow persists. Third, since there is no diffraction, there is consequently no mechanism to generate a circumferential wave.

Figure 8 shows frames from a movie depicting the scattering of an impulsive plane pressure wave (peak value=1) from a pressure release sphere, calculated using the raypath scatter method. In this case, to fulfill the boundary conditions, the field reflected from the sphere surface must be negative, to force the field directly at the boundary to zero. The animation was produced using the same method outlined for Fig. 7, but plotting the total field pressure as $p_0 - p_{23}$. When Fig. 8 is compared with Fig. 4, we again see that the expanding radial distribution and amplitude decay of the geometric wave front are practically identical in both cases. However, the effects of diffraction are again absent. The negative region behind the expanding wave front, and the filling of the acoustic shadow, are not observed in Fig. 8.

While the differences between Figs. 7 and 8 and the predictions of the Anderson theory are notable, the results of the two approaches are similar enough to suggest that, outside of the forward scattering region, raypath methods may be sufficiently accurate for many object scattering applications. This is indicated in Fig. 1(a). In the rigid sphere case (solid black line), the backscattered impulse response contains the region of negative pressure immediately behind the geometric spike and, later on, the positive circumferential wave peak, but these are both dominated by the amplitude of the geometric spike. Also, the observation of Figs. 2 and 4 suggests that this will continue to be true for a wide range of azimuthal angles on either side of the backscattering axis, except within, or close to, the forward scattering cone. In the pressure release case, the negative geometric backscattering spike is essentially the only feature of importance in the impulse response. Overall, a comparison of Figs. 2 and 4 with Figs. 7 and 8 suggests that simple raypath constructions can be used most of the time for impulsive (i.e., very broadband) incident waves in backscattering applications.

D. Animation produced using the Hickling and Wang theory

The essential difference between the theory of Hickling and Wang[11] and that of Anderson occurs in the scattering coefficient for the dipole term. It is the dipole component of the incident field that induces the translational motion of the sphere, and this must be properly incorporated into the theory in order to prescribe the physical behavior. It is very helpful for teaching purposes to include an animation for a movable rigid sphere, since the student is enabled to see the

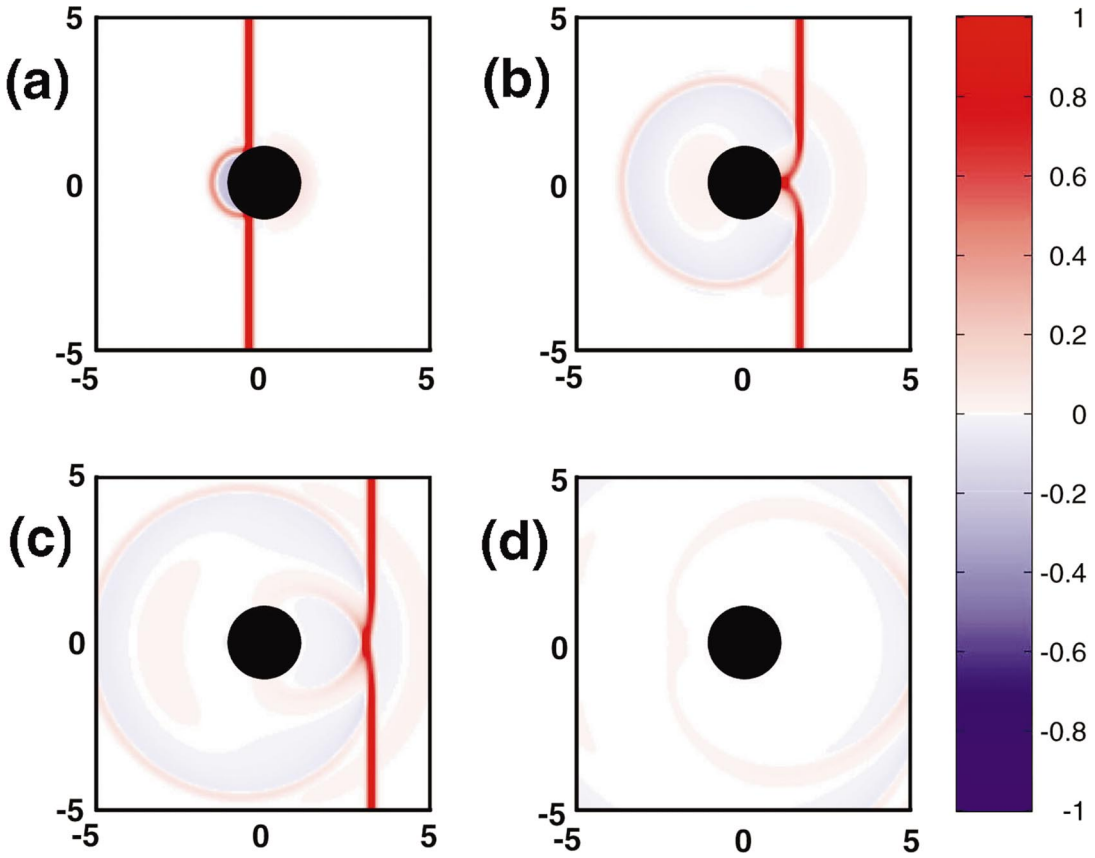


FIG. 9. Impulse scattering from a movable rigid sphere. The pressure field is plotted, and scaled to the peak value (=1) of the incident field. The scatterer is represented by a black disk, since no field is generated inside the sphere. The axes are scaled in radii R. The density of the sphere is 6% greater than the external fluid. This figure was produced using the Hickling and Wang scattering theory.

dipolar reaction of the object to the incident field, as evidenced by the compression of the acoustic field as the sphere ''pushes'' the water in front of it. Later, the sphere is also seen to reverse its direction and compress the water in the backward direction.

The Hickling and Wang scattering theory for a rigid sphere of finite density (i.e., a ''movable'' rigid sphere) applies rigid boundary conditions at the edges of the sphere, and allows the scatterer to move, or ''rebound,'' in response to the incident field without changing its spherical shape. The theory is not reproduced here, and the reader is referred to the Hickling and Wang paper.[11] The modification of the boundary conditions to allow for the motion of the sphere gives rise to the Hickling and Wang prescription for calculating the scattered field. This incorporates a change to the expression for the coefficient A_m appearing in Eq. (2), i.e.,

$$A_m = -i^m(2m+1)\begin{cases} \dfrac{g\ kR\ j'_m(kR) - j_m(kR)}{g\ kR\ h'_m(kR) - h_m(kR)} & (m=1), \\ \dfrac{j'_m(kR)}{h'_m(kR)} & (m\neq 1), \end{cases} \tag{16}$$

which reduces to Eq. (2) as $g\to\infty$, or when kR becomes very large.

The scattered impulse response for a movable rigid sphere was determined by using the Hickling and Wang expression for A_m to compute the complex scattered pressure $P_s(f)$ [via Eq. (2)] for a range of frequencies, assuming a constant incident amplitude as a function of frequency, and then using the inverse discrete Fourier transform [via Eq. (6)] in exactly the same manner as described above for the Anderson theory examples. The number of points in the inverse FFT, and the frequency bandwidth, were identical to those used previously.

In Fig. 1(a), the gray line shows the farfield backscattered pressure impulse response for a movable rigid sphere whose density is 6% higher than the surrounding fluid. The overall response is similar to that for the fixed rigid sphere (solid black line), except that the positive and negative excursions of the pressure field are more accentuated in the movable sphere case as a result of the rebound motion of the scatterer. The line shows that the geometric scattering spike is again followed by a negative pressure region, which, in this case, later reverses to become a positive pressure region, before ending with the circumferential wave peak.

Figure 9 shows movie frames depicting the scattering for the movable rigid sphere represented by the gray line in Fig. 1(a). The pressure field is plotted, and scaled throughout to the peak value (=1) of the incident field. Evidence of the rebound motion is seen in Figs. 9(a)–9(c). A broad ridge of expanding high pressure on the back side of the sphere (which is initiated *before* the arrival of the incident field at this location) may be clearly observed. This is due to the

sphere moving in response to the incident field and compressing, or ''pushing,'' the water in front of it. It is useful to compare Figs. 9(a)–9(c) with Figs. 2(a)–2(c). In Fig. 2, no corresponding high-pressure ridge may be seen, because the huge mass of the scatterer in the fixed rigid sphere case inhibits the rebound motion. In the movable sphere case, after the incident field has passed, the sphere also reverses direction and compresses the water adjacent to its front face, giving rise to a broader high-pressure wave in the backward direction [see Figs. 9(b), 9(c)]. Later, as the incident wave passes, diffraction into the shadow behind the sphere initiates a circumferential wave [see Figs. 9(c), 9(d)], which propagates around and breaks off in the backward direction [cf. Figs. 2(c), 2(d)].

III. CONCLUSIONS

The educational value of time–domain animations for visualizing acoustic impulse scattering from a sphere has been demonstrated. Modern computational methods have been used to accurately depict the complete ''cause and effect'' sequence of events that takes place during a complex physical interaction, and to enable important insights to be gained into the physical processes that lead to the formation of scattered wave fronts.

Anderson's fluid sphere theory was used to create animations that demonstrate scattering for two cases: (a) a fixed rigid sphere; (b) a pressure release sphere. By calculating the impulse response for a range of locations around the sphere, movies were produced that depict the individual wave processes comprised in the scattering behavior, and enable the construction of scattered wave fronts to be visualized. The backscattering regime was shown to be dominated by geometric reflections from the face of the sphere. The forward scatter region was examined, and the process of wave front healing directly observed. Results for both the fixed rigid and pressure release cases show that the incident and forward scattered fields combine to rapidly minimize amplitude and phase perturbations of the forward propagating wave front caused by the sphere. Diffraction into the acoustic shadow behind the sphere is visible and, in the rigid case, leads to a clearly discernible circumferential wave that eventually breaks off and propagates in the backward direction.

The raypath scatter method was used to create animations for fixed rigid and pressure release spheres. In both cases, geometrical scatter from the surface of the sphere is depicted, with an expanding radial distribution and amplitude decay of the wave front equivalent to the corresponding Anderson theory calculations. Features due to the effects of diffraction are absent, particularly diffraction from the incident wave behind the sphere to fill the acoustic shadow and heal the wave front. For many applications that require just the backscattered wave component, the raypath scatter method appears to be a very reasonable approximation.

The Hickling and Wang theory was used to create an animation to demonstrate scattering for a movable rigid sphere. Much of the response is similar to the fixed rigid sphere, but the variations of the pressure field are more accentuated due to the rebound motion of the scatterer. A broad ridge of expanding high pressure on the back side of the sphere is initiated before the arrival of the incident field, due to the sphere moving in response to the incident field and compressing the water in front of it.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (PE 61153N). The author would like to thank C. S. Clay for suggesting the comparison of the Anderson movies with the geometric raypath animations, and for outlining the method for calculating the geometric impulse response.

[1] L. Bragg and J. F. Nye, ''A dynamical model of a crystal structure,'' Proc. R. Soc. London, Ser. A **190**, 474–481 (1947); L. Bragg and W. M. Lomer, ''A dynamical model of a crystal structure II,'' *ibid.* **196**, 171–181 (1948).
[2] W. G. Neubauer, ''Experimental observation of three types of pulsed circumferential waves on solid aluminum cylinders,'' J. Acoust. Soc. Am. **44**, 1150–1152 (1968); ''Pulsed circumferential waves on aluminum cylinders in water,'' *ibid.* **45**, 1134–1144 (1968).
[3] W. G. Neubauer and L. R. Dragonette, ''Observation of waves radiated from circular cylinders caused by an incident pulse,'' J. Acoust. Soc. Am. **48**, 1135–1149 (1970); ''A schlieren system used for making movies of sound waves,'' *ibid.* **49**, 410–411 (1971).
[4] R. Hickling, ''Computer visualizations of the scattering of sound by structures in water,'' Acoust. Phys. **40**, 453–454 (1994) (also see list of references).
[5] V. W. Sparrow and D. A. Russell, ''Animations created in *MATHEMATICA* for acoustics education,'' J. Acoust. Soc. Am. **103**, 2987 (1998).
[6] D. A. Russell, ''Animations for teaching more advanced acoustics topics,'' J. Acoust. Soc. Am. **106**, 2197 (1999).
[7] D. A. Russell, ''Acoustics and vibration animations,'' http://www.gmi.edu/~drussell/demos.html. This useful website contains animations that visualize numerous concepts in acoustics and vibration.
[8] G. C. Gaunaurd and H. C. Strifors, ''Transient resonance scattering and target identification,'' Appl. Mech. Rev. **50**, 131–148 (1997).
[9] V. C. Anderson, ''Sound scattering from a fluid sphere,'' J. Acoust. Soc. Am. **22**, 426–431 (1950).
[10] H. Medwin and C. S. Clay, *Fundamentals of Acoustical Oceanography* (Academic, San Diego, 1998).
[11] R. Hickling and N. M. Wang, ''Scattering of sound by a rigid movable sphere,'' J. Acoust. Soc. Am. **39**, 276–279 (1966).
[12] S. Stein and M. Wysession, *An Introduction to Seismology, Earthquakes, and Earth Structure* (Blackwell, Malden MA, 2003), Sect. 2.5 (note particularly Fig. 2.5-19).
[13] G. Nolet and F. A. Dahlen, ''Wave front healing and the evolution of seismic delay times,'' J. Geophys. Res. **105**, 19 043–19 054 (2000).
[14] See EPAPS Document No. EPAPS-JASMAN-115-711404 for downloadable files containing the *QuickTime* movie animations. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps. See the EPAPS homepage for more information.
[15] N. J. Higham and D. J. Higham, *MATLAB Guide* (SIAM Press, Philadelphia, 2000).
[16] V. W. Sparrow, J. L. Rochat, and B. A. Bard, ''The use of color in the scientific visualization of acoustical phenomena,'' J. Comput. Acoust. **4**, 203–223 (1996).
[17] C. Feuillade and C. S. Clay, ''Anderson (1950) revisited,'' J. Acoust. Soc. Am. **106**, 553–564 (1999).
[18] R. K. Johnson, ''Sound scattering from a fluid sphere revisited,'' J. Acoust. Soc. Am. **61**, 375–377 (1977).
[19] R. Hickling, ''Analysis of echoes from a solid elastic sphere in water,'' J. Acoust. Soc. Am. **34**, 1582–1592 (1962).
[20] C. Feuillade, R. W. Meredith, N. P. Chotiros, and C. S. Clay, ''Time domain investigation of transceiver functions using a known reference target,'' J. Acoust. Soc. Am. **112**, 2702–2712 (2002).
[21] H. Uberall, R. D. Doolittle, and J. V. McNicholas, ''Use of sound pulses for a study of circumferential waves,'' J. Acoust. Soc. Am. **39**, 564–578 (1966).
[22] H. Uberall, ''Surface waves in acoustics,'' in *Physical Acoustics, Principles and Methods, Vol. X*, edited by W. P. Mason and R. N. Thurston (Academic, New York, 1973), Chap. 1.

Analysis of transient Lamb waves generated by dynamic surface sources in thin composite plates

Sauvik Banerjee and Ajit K. Mal[a)]
Mechanical and Aerospace Engineering Department, University of California, Los Angeles, California 90095-1597

William H. Prosser
Nondestructive Evaluation Branch, NASA Langley Research Center, MS 231, Hampton, Virginia 23681-0001

(Received 15 October 2003; revised 3 February 2004; accepted 10 February 2004)

A theoretical analysis is carried out in an effort to understand certain unusual properties of transient guided waves produced in a thin unidirectional graphite/epoxy composite plate by a localized dynamic surface load. The surface motion is calculated using an approximate plate theory, called the shear deformation plate theory (SDPT), as well as a recently developed finite element analysis (FEA), for their mutual verification. The results obtained by the two methods are shown to have excellent agreement. An interesting, nearly periodic ''phase reversal'' of the signal with propagation distance is observed for each propagation direction relative to the fiber direction. For clarification, a closed form analytical expression for the vertical surface displacement in an aluminum plate to an impulsive point force is obtained using the steepest descent method. It is found that the strong dispersion of the first antisymmetric waves at low frequencies is the main reason behind the phase reversal. This is verified further by measuring the surface response of a relatively thick aluminum plate to a pencil lead break source. The understanding developed in the paper is expected to be helpful in detecting and characterizing the occurrence of damage in composite structures. [DOI: 10.1121/1.1694993]

PACS numbers: 43.20.Bi, 43.35.Cg, 43.40.Le [ANN] Pages: 1905–1911

I. INTRODUCTION

Advanced composites are being used increasingly in aircraft, aerospace, marine, automotive and other structures, due to their high strength to weight ratio, formability and other favorable properties. However, these materials are highly sensitive to the presence of manufacturing and service related defects. As an example, graphite/epoxy laminate, one of the most widely used composite structural components, can sustain hidden internal damage (delamination, fiber breakage, matrix cracking) when it is subjected to foreign object impact. If undetected, the damage can grow, leading to catastrophic failure of the structure. Detection and characterization of hidden flaws in composites are critical factors for continued safe operation of advanced defects critical structures.

Ultrasonic nondestructive evaluation (NDE) techniques offer an efficient and accurate procedure for damage monitoring in composite structures. In order to develop reliable damage monitoring systems, it is necessary to have a clear understanding of the quantitative nature of ultrasonic waves that can be transmitted in composite laminates.

Propagation of elastic waves in isotropic solids has been investigated in great detail for several decades due to its importance in seismology[1] and ultrasonic nondestructive evaluation.[2] Waves in beams and plates have also been the subject of a large number of studies[3] since they are widely used in a variety of engineered structures. In contrast, wave propagation studies in structural composites are of relatively recent origin.[4,5] Since the composite laminates used in aircraft and aerospace structures are usually thin, approximate thin-plate theories have been developed in an effort to gain model-based understanding of the nature of the guided waves that can be transmitted in them.

The more important and relevant class of problems where the waves are generated by surface or subsurface sources is less well studied and is a topic of considerable current research interest. The solution of three-dimensional problems consisting of multilayered, angle-ply laminates of finite thickness and large lateral dimensions subjected to various types of surface loads has been given in Refs. 6–8. The finite element method (FEM) has also been used as an alternative to solve these problems. A dynamic finite element code has been developed by NIST for the prediction of AE wave propagation in plates.[9] This code has been validated with both experimental measurements and analytical predictions for a variety of source conditions and plate dimensions in isotropic materials.[9–11] However, these so-called ''exact'' methods are computationally intensive and cannot generally be used in damage monitoring systems, where data collected by on-board sensors must be analyzed in real-time.

As indicated earlier, in thin-walled aircraft and aerospace structures, approximate solutions based on thin plate theories can provide reasonably accurate numerical results with negligible computational effort. It has been shown that the approximate SDPT, where the transverse shear and rotary inertia are retained in modeling the dynamic deformations across the thickness of the plate, can provide accurate waveforms at lower frequencies (i.e., the plate thickness is small

[a)]Author to whom correspondence should be addressed. Electronic mail: ajit@ucla.edu

compared to the wavelength) and at some distance away from the source.[12] In this paper, the approximate method is used to calculate the response of a thin unidirectional composite plate to a concentrated force on the surface of the plate. The results from SDPT are compared with those from FEM for their mutual verification. A previously unknown feature of the surface motion, namely, a complete and spatially periodic reversal of the initial phase of the displacement pulse, is observed in the calculated waveforms using both methods. Since these methods are primarily numerical in nature, they do not provide a clear explanation of the reason behind this behavior of the waveforms. An approximate closed form analytical expression for the far field displacement produced by a surface line load on an isotropic plate is derived and is shown to have the same general feature. The analytical solution shows that the phase reversal is caused by the strong dispersion of the guided first antisymmetric or flexural waves in the plate.

II. NUMERICAL SOLUTIONS

The first order SDPT retaining transverse shear and rotary inertia of the plate elements is used here. Assuming that the x-y plane is the mid-plane of the laminate, the displacement components within the laminate are assumed to be of the form

$$u(x,y,z,t)=u_0(x,y,t)+z\psi_x(x,y,t),$$
$$v(x,y,z,t)=v_0(x,y,t)+z\psi_y(x,y,t),$$
$$w(x,y,z,t)=w_0(x,y,t),$$

where (u_0,v_0,w_0) are the displacement components at a point in the mid-plane, and ψ_x and ψ_y are the rotations of a line element, originally perpendicular to the longitudinal plane, about the y and x axes, respectively. The solution of the resulting approximate system of equations can be obtained using multiple integral transforms. A detailed formulation on first order SDPT for wave field calculations in composite laminates can be found in Ref. 12 and the associated literature cited therein, and will not be repeated here. In general, the Fourier time transform of the displacement and stress components on the surface of the laminate can be expressed as wavenumber integrals in the form

$$F(\omega)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(k_1,k_2,\omega)e^{i(k_1x_1+k_2x_2)}dk_1dk_2,$$

where ω is the circular frequency, $F(\omega)$ is the Fourier time transform of the source function, $f(t)$, x_1, x_2 are the coordinates of the field point on the surface of the plate with the x_1-axis directed along the fibers, and the function g is determined from the equations of motion and the boundary conditions of the problem. These integrals must, in general, be evaluated numerically for a range of frequencies and the resulting spectra can then be inverted by FFT to determine the time dependent displacement and stress components. The computational effort required to evaluate the integrals numerically accurately is quite large. An adaptive integration scheme developed in Ref. 12 is used to calculate the surface motion. It should be noted that the result includes both the body and guided wave contributions, however, at lower frequencies most of the energy in the plate is carried by the first antisymmetric guided wave or the A0 mode.

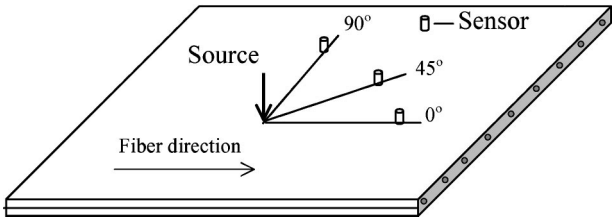


FIG. 1. Schematic of a loaded unidirectional composite plate showing position of the sensors wrt the fiber direction.

The normal surface displacement in a unidirectional graphite/epoxy composite plate to a normal point load on the surface calculated from the SDPT and FEM is presented first. A loaded unidirectional composite plate with surface mounted sensors is shown in Fig. 1. The elastic properties of the graphite/epoxy composite material used in the calculations are given in Table I.

The plate thickness is taken as 1 mm and the load is assumed to be of the form

$$F(x_1,x_2,t)=f(t)g(x_1,x_2), \tag{1}$$

where

$$f(t)=\sin\left(\frac{2\pi t}{\tau}\right)-0.5\sin\left(\frac{4\pi t}{\tau}\right), \quad 0<t<\tau,$$
$$=0, \quad t>\tau. \tag{2}$$

For a point load,

$$g(x_1,x_2)=\delta(x_1)\,\delta(x_2), \tag{3a}$$

and for a load uniformly distributed in a circle of radius a,

$$g(x_1,x_2)=1, \ x_1^2+x_2^2\leqslant a^2$$
$$=0, \ x_1^2+x_2^2\geqslant a^2. \tag{3b}$$

The time dependence of the load $f(t)$ and its Fourier transform are plotted for $\tau=1$ μs in Fig. 2. It can be seen that the source spectrum is maximum around 1.1 MHz and becomes negligibly small beyond 3 MHz.

The results obtained from the SDPT for the vertical surface displacement at a number of field points are compared with those from FEM in Figs. 3(a)–(c) for propagation along 0°, 45°, and 90° with respect to the fiber direction. A third-order elliptic digital filter (Matlab 6.1) with a pass band of 0.05–0.65 MHz is applied to all calculated spectra to eliminate high frequency numerical noise. It can be seen that the agreement between the results from the two methods is excellent. Since the approximate theory (SDPT) overestimates the flexural wave speed,[12] the arrival time of waves in this case is somewhat earlier than that in the FEM.

TABLE I. Material constants of graphite/epoxy composite material.

Density (g/cm^3)	C_{11} (GPa)	C_{12} (GPa)	C_{22} (GPa)	C_{23} (GPa)	C_{55} (GPa)
1.578	160.73	6.44	13.92	6.92	7.07

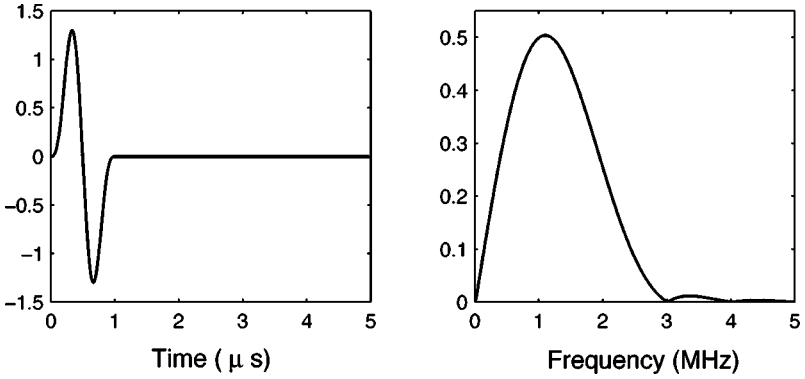
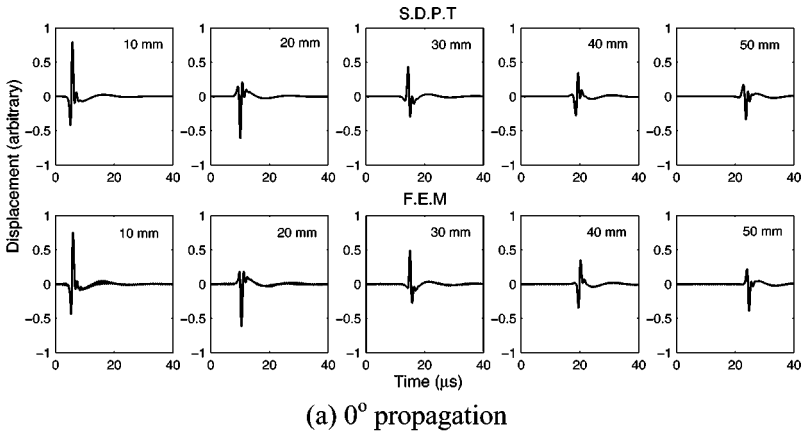
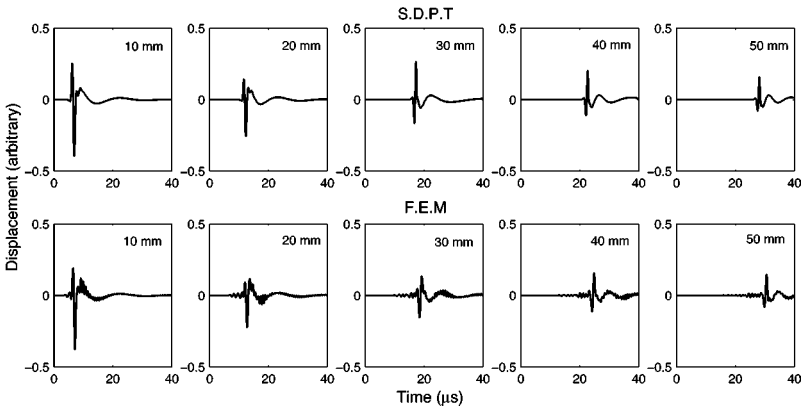


FIG. 2. Time history and spectrum of the source in arbitrary units.

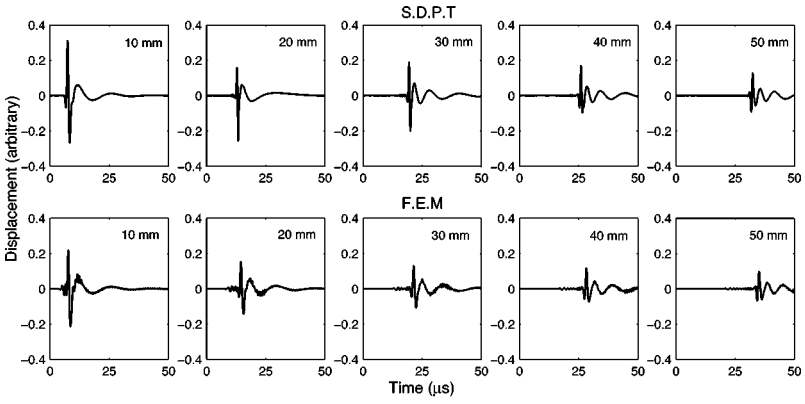


(a) 0° propagation



(b) 45° propagation

FIG. 3. Time history of vertical surface displacement in a 1-mm-thick unidirectional graphite/epoxy composite plate subjected to a point load from SDPT (first row) and FEM (second row).



(c) 90° propagation

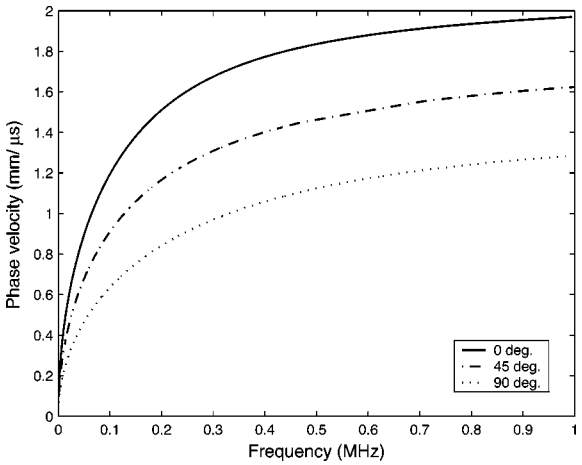


FIG. 4. Dispersion curves for unidirectional graphite/epoxy composite plate of 1-mm thickness for different propagation direction wrt the fibers calculated from SDPT.

The most noteworthy feature of the surface displacements calculated by both methods is a periodic reversal in the initial phase of the pulse with propagation distance. For propagation along 0° [Fig. 3(a)], the reversal occurs at an interval of 10 mm, whereas it occurs at larger intervals for propagation along 45° [Fig. 3(b)]; the interval is the largest for propagation along 90° [Fig. 3(c)]. This is due to faster flexural wave speed along 0° direction as shown in the dispersion curves for the plate for different propagation directions (Fig. 4). Figure 5 shows the response at distances between 10 and 20 mm at an interval of 2 mm for propagation along 0°. It can be seen that the rising first peak and decaying second peak result in a smooth reversal of phase at an interval of 10 mm.

The time history of the surface displacement due to a uniform distributed surface load on a 2.5-mm-radius circular region is shown in Fig. 6 for propagation along 0° at different locations. The periodic reversal in the phase is found in this case, too. The waveforms do not change as significantly as in the point load case, and their sharpness is reduced considerably due to the spatial distribution of the load.

III. PHASE REVERSAL: THEORETICAL AND EXPERIEMTAL VALIDATION

The phase reversal of the surface displacement is difficult to explain from the primarily numerical methods used in both the SDPT and the FEM. A simpler problem for which an analytical solution can be obtained is considered here in an effort to understand the source of this behavior of the guided waves. The problem involves an isotropic (aluminum) plate of thickness $2H$ subjected to a normal concentrated line load. Considering only A0 mode, the Fourier transform of the normal surface displacement, V, can be expressed in the form

$$V(x,H,\omega)=\frac{F(\omega)}{4\pi\mu}\int_{-\infty}^{\infty}\frac{k_2^2\eta_1 e^{ikx}}{(2k^2-k_2^2)^2\tanh(\eta_1 H)-4k^2\eta_1\eta_2\tanh(\eta_2 H)}dk, \tag{4}$$

where

$$\eta_j=\sqrt{k^2-k_j^2};\quad k_j=\frac{\omega}{c_j},\quad j=1,2,$$

c_1 and c_2 are the longitudinal and shear wave speeds in the material of the plate, $F(\omega)$ is the Fourier time transform of the load and μ is the shear modulus.

At low frequencies, when the plate thickness is small compared to the wavelength, the quantities kH, k_1H, k_2H are small. Then, retaining the first two terms in the Taylor series expansion of $\tanh(\eta_1 H)$ and $\tanh(\eta_2 H)$, Eq. (4) can be rewritten as

$$V(x,H,\omega)=\frac{F(\omega)}{4\pi\mu}\int_{-\infty}^{\infty}\frac{(k_2^2/H)e^{ikx}}{R(k)}dk, \tag{5}$$

where

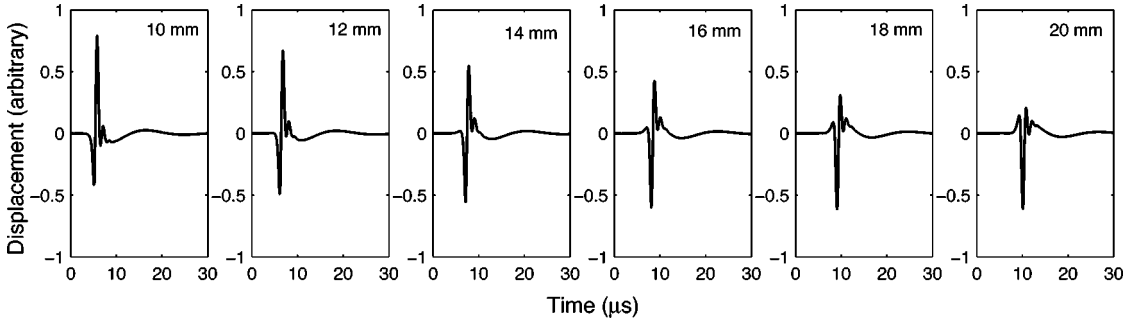


FIG. 5. Time history of vertical surface displacement in a 1-mm-thick unidirectional graphite/epoxy composite plate subjected to a point load at an interval of 2 mm in between 10 and 20 mm, showing smooth change in phase, for 0° propagation (SDPT).

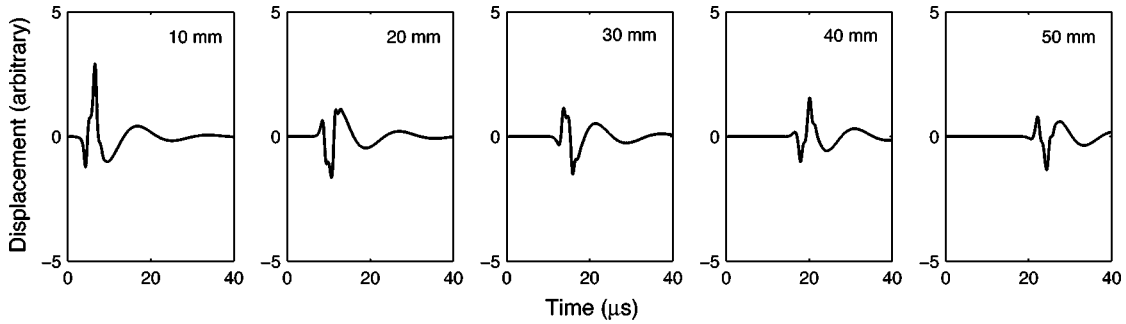


FIG. 6. Time history of vertical surface displacement in a 1-mm-thick unidirectional graphite/epoxy composite plate under uniformly distributed load on a circle of 2.5-mm radius, for propagation along 0° (SDPT).

$$R(k)=k_2^4-\frac{4}{3}k^4(k_2H)^2(1-\alpha^2)-k^2\frac{(k_2H)^2}{3}$$

$$\times k_2^2(4\alpha^2-3)+k_2^4\frac{(k_1H)^2}{3}+O[(k_1H)^3], \tag{6}$$

and $\alpha=k_1/k_2=c_2/c_1$. The last term in (6) is third order in k_1H and is neglected in further calculations.

Applying the contour integration technique in the complex k-plane, the propagating guided wave contribution to the surface displacement is given by the residue at the pole of the integrand as

$$V(x,H,\omega)\cong\frac{iF(\omega)}{2\mu}\frac{(k_2^2/H)e^{ik_px}}{R'(k_p)} \tag{7}$$

where k_p, the wavenumber of the guided waves in the plate, is a real root of $R(k)=0$. Assuming that the material of the plate is aluminum, $\alpha=\frac{1}{2}$,

$$R(k)\cong-\frac{k_2^2}{H^2}\left\{(kH)^4-\frac{2}{3}(kH)^2(k_2H)^2-(k_2H)^2\right.$$

$$\left.\times\left(1+\frac{(k_2H)^2}{12}\right)\right\} \tag{8}$$

and

$$R'(k)\cong-\frac{4k_2^2}{H}(kH)\left[(kH)^2-\frac{1}{3}(k_2H)^2\right]. \tag{9}$$

Then the root of $R(k)$ is given by the approximate expression

$$k_pH\cong\sqrt{k_2H}[1+\tfrac{1}{6}(k_2H)] \tag{10}$$

or

$$k_2H\cong(k_pH)^2[1-\tfrac{1}{3}(k_pH)^2]. \tag{11}$$

The time dependent displacement can be written in the alternate integral form as

$$v(x,H,t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}V(x,H,\omega)e^{-i\omega t}\,d\omega$$

$$=\int_{-\infty}^{\infty}G(k)e^{i[kx-\omega(k)t]}dk, \tag{12}$$

where

$$G(k)=\frac{1}{2\pi}V(x,H,\omega)\frac{d\omega}{dk}. \tag{13}$$

In general, the integral in (12) must be evaluated numerically to obtain the pulse shape of the traveling waves. We use the steepest descent method to derive an approximate analytical expression for the integral. The stationary values of the exponent $g(k)=kx-\omega(k)t$ occurs near the values of k given by

$$\frac{d\omega}{dk}=\frac{x}{t}=U(k), \tag{14}$$

where $U(k)$ is the group velocity of the plate guided waves. From Eq. (11),

$$\frac{\omega H}{c_2}=(kH)^2-\frac{1}{3}(kH)^4 \tag{15}$$

so that

$$\frac{d\omega}{dk}=2c_2kH\left[1-\frac{2}{3}(kH)^2\right]=\frac{x}{t}. \tag{16}$$

Let the solution of (14) and (16) be $k=k_o$ and $\omega_o=\omega(k_o)$. Then

$$k_oH\cong\frac{x}{2c_2t}+\frac{2}{3}\left(\frac{x}{2c_2t}\right)^3 \tag{17}$$

and

$$\omega_o\cong\frac{c_2}{H}\left[\left(\frac{x}{2c_2t}\right)^2+\left(\frac{x}{2c_2t}\right)^4\right]. \tag{18}$$

Therefore,

$$k_ox-\omega_ot\cong\frac{x^2}{4c_2tH}+\frac{x^4}{48(c_2t)^3H}. \tag{19}$$

Expanding $g(k)$ in a Taylor series about $k=k_o$ and, after some manipulation, the integral (12) gives

$$v(x,H,t)=\left\{\left[\frac{2\pi}{|-g''(k_o)|}\right]^{1/2}G(k_o)e^{i[g(k_o)-\pi/4]}\right\}. \tag{20}$$

Retaining second order terms in the nondimensional quantity, $x/2c_2t$, the analytical expression for the normalized surface displacement, $\overline{v}$, due to a source of time dependence $f(t)=S_o\delta(t)$ in an aluminum plate can be written explicitly as

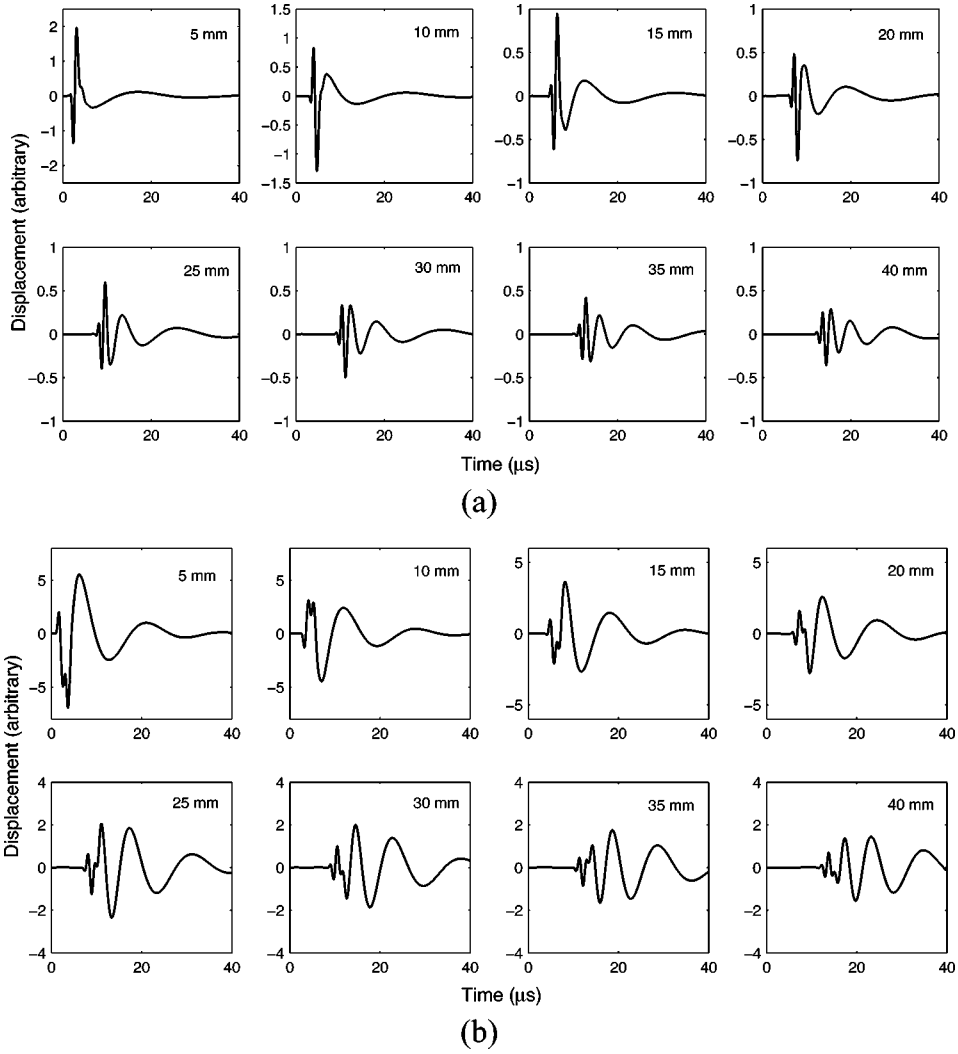


FIG. 7. Time history of vertical surface displacement in a 1-mm-thick aluminum plate subjected to (a) a point load and (b) a uniformly distributed load on a 2.5-mm radius (SDPT).

$$\bar{v} = \frac{v(x,H,t)}{S_o c_2/8\pi\mu H} \cong \sqrt{\frac{\pi H}{c_2 t}} \frac{1-\frac{2}{3}(x/2c_2t)^2}{(x/2c_2t)^2\sqrt{|1-2(x/2c_2t)^2|}} \sin\left[\frac{x^2}{4c_2Ht} - \frac{\pi}{4}\right]. \tag{21}$$

It can be seen that expression (21) carries a singularity that travels with a speed $V_s = \sqrt{2}c_2$, and that

$$\bar{v} > 0, \text{ for } 2\sqrt{2}\left(2n+\frac{1}{4}\right)\pi < \frac{x}{H} < 2\sqrt{2}\left(2n+\frac{5}{4}\right)\pi,$$

$$\bar{v} < 0, \text{ for } 2\sqrt{2}\left(2n+\frac{5}{4}\right)\pi < \frac{x}{H} < 2\sqrt{2}\left(2n+\frac{9}{4}\right)\pi,$$

where n is a positive integer. Thus the distance required for a complete phase reversal of this singularity in the waveform is given by $2\sqrt{2}\pi H$.

The response of the aluminum plate to the normal surface load is calculated using SDPT. The vertical surface displacement produced in an aluminum plate of 1-mm thickness by the point and distributed loads used earlier in the composite plate case are plotted in Fig. 7 at an interval of 5 mm up to 40 mm. The same band pass filter as that of composite plate case is applied to the spectra before inversion into time domain. The numerical results show that the surface response reverses its phase at an interval of approximately 5 mm, or $2\sqrt{2}\pi H$ as predicted by the approximate analytical solution. However, this effect becomes negligible at large distances from the source.

The calculated and measured signals for the Hsu–Neilsen-type pencil lead break source[13,14] are also shown in Fig. 8 for an aluminum plate of thickness 3.1. The detail of the experimental setup[14] is omitted here for brevity. Both calculated and experimentally acquired waveforms are normalized relative to their individual peak amplitudes. The resulting acoustic emission plots show good agreement for every separation distance between the lead break location and the recording transducer. Note that after the first few wiggles in acoustic emission time-history, reflections are seen in experimentally acquired waveforms. Since the experimentally acquired waveform is triggered by a constant triggering transducer at a fixed distance away from the source and is not plotted on an absolute time scale, only the waveforms and the relative times between the two plots can be compared. More notably, phase reversal is seen at an interval of

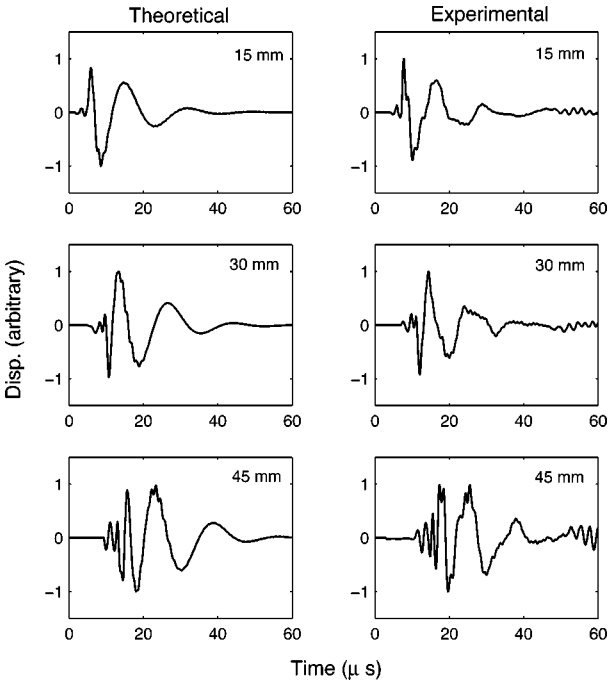


FIG. 8. Time history of vertical surface displacement in a 3.1-mm-thick aluminum plate to a pencil lead break source.

approximately 15 mm($\approx 2\sqrt{2}\pi H$), which was deliberately chosen as a recording interval for the test.

IV. CONCLUDING REMARKS

The good agreement between the results from SDPT and FEM indicates that both methods work well for the calculation of the wave field produced in a plate by localized dynamic sources. Since the SDPT is expected to be more accurate at larger distances from the source, while the FEM is more efficient at smaller distances, the two methods can be combined for efficient and accurate calculation of the waveforms in plates. This capability is expected to be helpful in the accurate characterization of AE sources, and to enhance the accuracy of data interpretation for online monitoring of impact damage critical structural components.

Although the closed form analytical expression (21) differs significantly from the exact solution, the usefulness of the approximate solution to understand the physics of the problem cannot be overemphasized. The approximate solution clearly shows that the dispersive nature of waves is the source of the unexpected phase reversal of the waveforms during propagation on the surface of the plate. It should be recalled that the low frequency symmetric or extensional waves do not have this property, nor do nondispersive Rayleigh waves propagating in a half-space.

ACKNOWLEDGMENT

This research was supported by NASA Langley Research Center under Grant No. NAG-1-02036.

[1] K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods* (Freeman, San Francisco, 1980).
[2] J. D. Achenbach, *Wave Propagation in Elastic Solids* (North-Holland, Amsterdam, 1973).
[3] J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge U.P., New York, 1999).
[4] A. H. Nayfeh, *Wave Propagation in Layered Anisotropic Media with Applications to Composites* (Elsevier, Amsterdam, 1995).
[5] D. E. Chimenti, ''Guided waves in plates and their use in material characterization,'' Appl. Mech. Rev. **50**, 247–284 (1997).
[6] A. K. Mal, ''Wave propagation in layered composite laminates under periodic surface loads,'' Wave Motion **10**, 257–266 (1988).
[7] A. K. Mal and S. S. Lih, ''Elastodynamic response of a unidirectional composite laminate to concentrated surface loads: parts I & II,'' J. Appl. Mech. **55**, 878–892 (1992).
[8] S. S. Lih and A. K. Mal, ''Response of multilayered composite laminates to dynamic surface loads,'' Composites, Part B **27B**, 633–641 (1996).
[9] W. H. Prosser, M. A. Hamstad, J. Gary, and A. O'Gallagher, ''Comparison of finite element and plate theory methods for modeling acoustic emission waveforms,'' J. Nondestruct. Eval. **18**(3), 83–90 (1999).
[10] J. Gary and M. A. Hamstad, ''On the far-field structure of waves generated by a pencil break on a thin plate,'' J. Acoust. Emiss. **12**(3-4), 157–170 (1994).
[11] D. Guo, A. K. Mal, and M. A. Hamstad, ''AE wavefield calculations in a plate,'' J. Acoust. Emiss. **16**(1-4), S222–S232 (1998).
[12] S. S. Lih and A. K. Mal, ''On the accuracy of approximate plate theories for wave field calculations in composite laminates,'' Wave Motion **21**, 17–34 (1995).
[13] F. Breckenridge, T. Proctor, N. Hsu, S. Fick, and D. Eitzen, ''Transient sources for acoustic emission work,'' in *Progress in Acoustic Emission V*, edited by K. Yamguchi, H. Takakashi, and H. Niitsuma (JSNDI, Tokyo, Japan, 1990), pp. 20–37.
[14] D. Guo, A. K. Mal, and K. Ono, ''Wave theory of acoustic emission in composite laminates,'' J. Acoust. Emiss. **14**, S19–S46 (1996).

Guided circumferential shear horizontal waves in an isotropic hollow cylinder

Xiaoliang Zhao[a] and Joseph L. Rose[b]
Department of Engineering Science & Mechanics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 30 March 2003; revised 31 October 2003; accepted 22 December 2003)

Guided time-harmonic shear horizontal (SH) waves propagating in the circumferential direction of an isotropic hollow cylinder are studied. The dispersion equation as well as the displacement and stress field across the wall thickness is derived analytically. Compared with the SH waves in a plate, a quantitative guideline of how well a plate model can approximate a pipe in the circumferential direction is given for defect characterization purpose. The work is also crucial for initiating work efforts on three-dimensional wave scattering for pipeline inspection. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1691037]

PACS numbers: 43.20.Bi, 43.20.Mv [SFW] Pages: 1912–1916

I. INTRODUCTION

Axial cracks and corrosion defects are often found in both industrial and military hollow cylindrical structures like pipelines and cylindrical containers. Reliable and easy-to-use inspection systems are in great need to locate the defects and to be able to characterize and size them efficiently. Hirao and Ogi[1] proposed a circumferential SH-wave Electromagnetic Acoustic Transducer (EMAT) technique for detecting corrosion defects on the outer surface of steel pipelines with and without protective resin coating. Gauthier[2] used multimode SH waves generated by EMATs to form B-scan images of a defect on a pipe. The reflection and transmission coefficients of SH waves passing through a two-dimensional surface-breaking defect or a stringer-like internal inclusion in a pipe was reported by the present authors.[3] However, all of these publications were based on an empirical plate-model approximation for a pipe of large diameter-to-wall-thickness ratio. A rigorous theory of guided SH waves propagating in the circumferential direction of a hollow cylinder needs to be established.[4]

Gazis[5,6] theoretically investigated guided waves that propagate in the axial direction and are resonant in the circumferential direction of a hollow cylinder. The case when the axial wave number is zero decouples into axially motion-independent plane–strain vibration and longitudinal shear vibration, both of which are standing waves in the circumferential direction. Liu and Qu[7] developed the model of guided plain–strain waves propagating in the circumferential direction of a hollow cylinder, with emphasis on the dispersion relation and displacement profile derivation and discussion. In this paper, the guided time-harmonic SH wave propagating in the circumferential direction of a hollow cylinder is studied. The dispersion equation as well as the displacement and stress distribution across the wall of the hollow cylinder is derived analytically. They are compared with that of the SH waves in a plate of the same thickness numerically. Relative errors of the phase and group velocities when using a plate model to approximate a pipe are given as a quantitative measure of whether that approximation is valid. With the approximated plate model, the guided wave interaction and scattering from a three-dimensional defect in a hollow cylindrical structure of large diameter-to-wall-thickness ratio could be tackled without extensive eigenmode and wave structure calculations in each scattering direction.[8]

II. THEORETICAL MODEL

Consider steady-state time-harmonic waves propagating in the circumferential direction of a hollow cylinder of inner radius a and outer radius b, as shown in Fig. 1. Assume the material is linearly elastic and isotropic, and the wave motion is independent of z. Two types of guided waves are possible in this hollow cylinder: one is the plane–strain vibration wave similar to the Lamb wave in a plate;[7] the other is the longitudinal shear wave, which will be studied in detail in this paper.

The analytical derivation of the frequency equation of the guided circumferential SH waves begins with Navier's equation of motion as Eq. (1), where $\mathbf{u}$ is the displacement vector, ρ is the density, λ and μ are Lamé's constants, respectively

$$\mu\nabla^2\mathbf{u}+(\lambda+\mu)\nabla\nabla\cdot\mathbf{u}=\rho(\partial^2\mathbf{u}/\partial t^2). \tag{1}$$

Consider only the displacement in the z direction in the cylindrical coordinates, i.e., $u_r=u_\theta=0$ and $u_z\neq 0$; Eq. (1) can be written as

$$\rho\frac{\partial^2 u_z}{\partial t^2}=\mu\left(\frac{\partial^2 u_z}{\partial r^2}+\frac{1}{r}\frac{\partial u_z}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u_z}{\partial\theta^2}\right). \tag{2}$$

Note that u_z is a function of r, θ, and t only. Since the wave is time harmonic and propagates along the θ direction, let

$$u_z=\Psi(r)e^{i(kb\theta-\omega t)},\quad a\leqslant r\leqslant b, \tag{3}$$

in which the wave number k is defined by $k=\omega/c(b)$. Here, $c(r)$ is the linear phase velocity for material particles located at distance r from the axis of the hollow cylinder. Equation (2) can be transformed into

[a]Presently with Intelligent Automation, Inc. Rockville, MD 20855. Electronic mail: xzhao@i-a-i.com
[b]Electronic mail: jlresm@engr.psu.edu

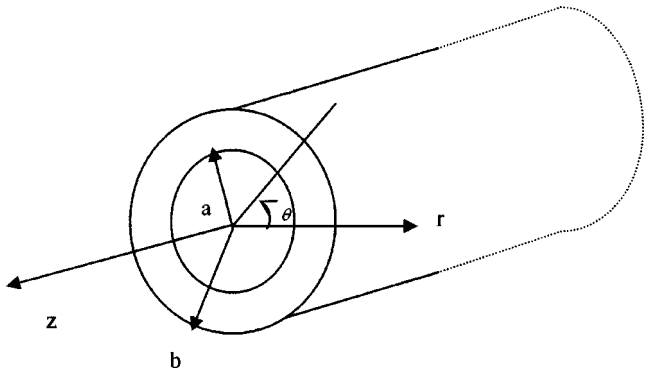


FIG. 1. Cylindrical coordinates of the hollow cylinder and dimensions.

$$\Psi'' + \frac{1}{r}\Psi' + \left[\frac{\omega^2}{c_T^2} - \left(\frac{kb}{r}\right)^2\right]\Psi = 0, \tag{4}$$

Here, $'$ and $''$ denotes the first- and second-order derivative with respect to r, respectively; $c_T = \sqrt{\mu/\rho}$ is the shear wave velocity. This equation is easy to recognize as a Bessel equation. Follow the convention used in Refs. 7 by introducing the nondimensional variables

$$\overline{r} = \frac{r}{b}, \quad \overline{k} = kh, \quad \overline{\omega} = \frac{\omega h}{c_T}, \quad h = b - a \tag{5}$$

$$\hat{k} = \frac{\overline{k}}{1-\eta}, \quad \hat{\omega} = \frac{\overline{\omega}}{1-\eta} = \frac{\omega b}{c_T}, \quad \eta = \frac{a}{b}. \tag{6}$$

The solutions can be written as

$$\Psi(r) = AJ_{\hat{k}}(\hat{\omega}\overline{r}) + BY_{\hat{k}}(\hat{\omega}\overline{r}), \quad a \leqslant r \leqslant b, \tag{7}$$

where $J_{\hat{k}}(x)$ and $Y_{\hat{k}}(x)$ are, respectively, the first and second kind of Bessel functions of order $\hat{k}$. A and B are arbitrary constants.

It is assumed that the surface of the hollow cylinder is traction free, i.e.,

$$\sigma_{rz}|_{r=a,b} = \mu \left.\frac{\partial u_z}{\partial r}\right|_{r=a,b} = 0. \tag{8}$$

Substituting Eqs. (3) and (6) into Eq. (8), we can get

$$AJ'_{\hat{k}}\left(\frac{\omega}{c_T}r\right) + BY'_{\hat{k}}\left(\frac{\omega}{c_T}r\right)|_{r=a,b} = 0, \tag{9}$$

or in the expanded form

$$\begin{cases} A\left[J_{\hat{k}-1}\left(\frac{\omega}{c_T}a\right) - J_{\hat{k}+1}\left(\frac{\omega}{c_T}a\right)\right] \\ \quad + B\left[Y_{\hat{k}-1}\left(\frac{\omega}{c_T}a\right) - Y_{\hat{k}+1}\left(\frac{\omega}{c_T}a\right)\right] = 0 \\ A\left[J_{\hat{k}-1}\left(\frac{\omega}{c_T}b\right) - J_{\hat{k}+1}\left(\frac{\omega}{c_T}b\right)\right] \\ \quad + B\left[Y_{\hat{k}-1}\left(\frac{\omega}{c_T}b\right) - Y_{\hat{k}+1}\left(\frac{\omega}{c_T}b\right)\right] = 0 \end{cases}. \tag{10}$$

This is a set of linear homogeneous algebraic equations with A and B the unknown variables. For nontrivial solutions, the determinant of the coefficient matrix of this system of equations must vanish. Thus, the dispersion relation between the wave number and the frequency of the circumferential guided SH wave in a hollow cylinder can be written as

$$\begin{aligned}&[J_{\hat{k}-1}(k_T a) - J_{\hat{k}+1}(k_T a)][Y_{\hat{k}-1}(k_T b) - Y_{\hat{k}+1}(k_T b)] \\ &\quad - [J_{\hat{k}-1}(k_T b) - J_{\hat{k}+1}(k_T b)][Y_{\hat{k}-1}(k_T a) \\ &\quad - Y_{\hat{k}+1}(k_T a)] = 0,\end{aligned} \tag{11}$$

where $k_T = \omega/c_T$ is the wave number of the shear wave. Since both $\hat{k}$ and k_T are frequency dependent, for a given frequency f (or k_T), $\hat{k}$ can be obtained from Eq. (11), and consequently the phase velocity $c(b) = 2\pi f/\hat{k}(1-\eta)$ and group velocity $c_g = d\omega/dk$. For each frequency point f, Eq. (11) may give many real roots, each of which corresponds to a possible wave mode that can propagate in the circumferential direction. Therefore, a family of phase velocity and group velocity versus frequency dispersion curves can be obtained. They provide a map of which kind of mode can be present in the structure and at what phase and group velocity, which is of great help in designing guided wave sensors and interpreting acquired inspection data. Note that in Gazis' article,[5] resonant modes in the circumferential direction of a hollow cylinder were discussed. They are special cases when $\hat{k}$ are integers.

Once $\hat{k}$ is obtained from Eq. (11), it can be substituted back to Eq. (10) so that the constant B can be expressed as a function of constant A. Thus, displacement amplitude $\Psi(r)$ can be expressed as

$$\begin{aligned}\Psi(r) = A\left[J_{\hat{k}}(k_T r) - \frac{J_{\hat{k}-1}(k_T a) - J_{\hat{k}+1}(k_T a)}{Y_{\hat{k}-1}(k_T a) - Y_{\hat{k}+1}(k_T a)} Y_{\hat{k}}(k_T r)\right], \\ a \leqslant r \leqslant b.\end{aligned} \tag{12}$$

Note that Eq. (12) actually describes the displacement amplitude variation with respect to r in the hollow cylinder wall. Once this displacement field u_z is obtained, the stress components of the circumferential SH waves can be calculated by Hooke's law. In the cylindrical coordinate system, the only nonzero stress components are

$$\begin{aligned}\tau_{rz} = \mu\frac{\partial u_z}{\partial r} = \frac{Ak_T\mu}{2}\Bigg\{&J_{\hat{k}-1}(k_T r) - J_{\hat{k}+1}(k_T r) \\ &- \frac{J_{\hat{k}-1}(k_T a) - J_{\hat{k}+1}(k_T a)}{Y_{\hat{k}-1}(k_T a) - Y_{\hat{k}+1}(k_T a)}[Y_{\hat{k}-1}(k_T r) \\ &- Y_{\hat{k}+1}(k_T r)]\Bigg\} e^{i(kb\theta - \omega t)}\end{aligned} \tag{13}$$

$$\begin{aligned}\tau_{\theta z} = \mu\frac{1}{r}\frac{\partial u_z}{\partial \theta} = \frac{ikbA\mu}{r}\Bigg[&J_{\hat{k}}(k_T r) \\ &- \frac{J_{\hat{k}-1}(k_T a) - J_{\hat{k}+1}(k_T a)}{Y_{\hat{k}-1}(k_T a) - Y_{\hat{k}+1}(k_T a)} Y_{\hat{k}}(k_T r)\Bigg] e^{i(kb\theta - \omega t)}.\end{aligned} \tag{14}$$

Note that both stress components vary along the r direction and propagate in the θ direction. They are an indispensable part of the elastic wave field in the hollow cylinder. One may notice that if we keep the fh value unchanged and increase η to 1, i.e., either both a and b go to ∞ while $b - a = h$ remains,

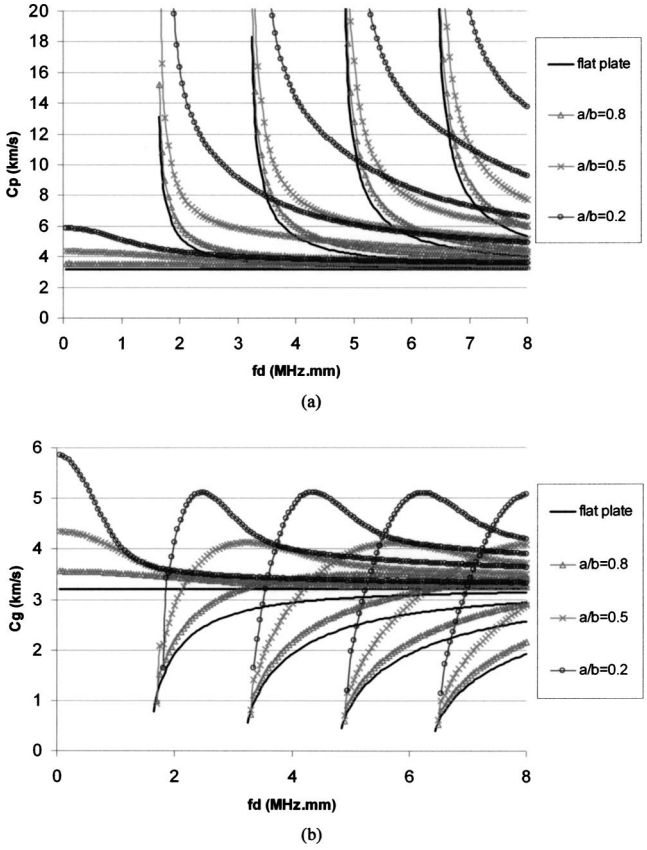


FIG. 2. (a) Phase velocity and (b) group velocity dispersion curve for SH wave in the circumferential direction of a hollow cylinder. In the legend, a is the inner radius, b is the outer radius of the cylinder.

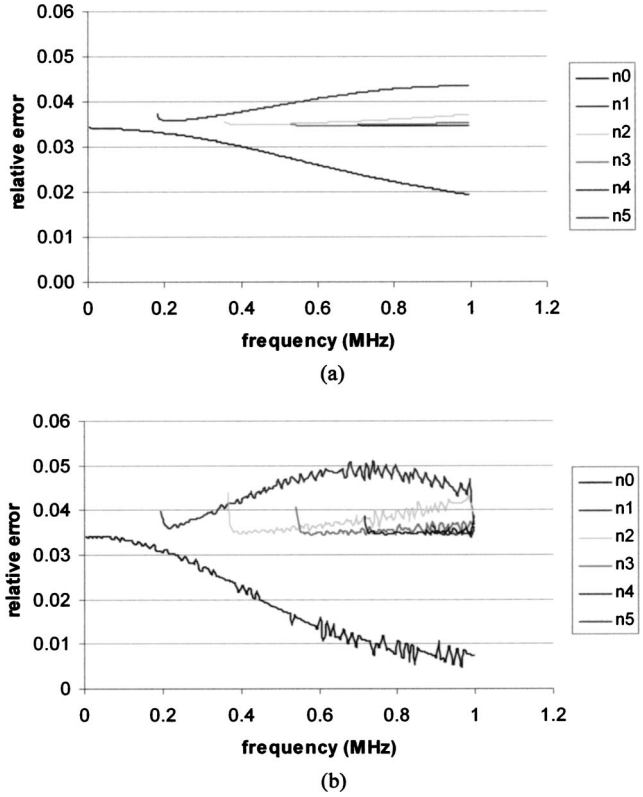


FIG. 3. (a) Phase velocity and (b) group velocity plate model approximation error for SH waves in the circumferential direction of a 10-in. schedule-40 pipe.

or $b-a\to 0$ while frequency $f\to\infty$, Eq. (4) will yield

$$\Psi''+\left[\frac{\omega^2}{c_T^2}-k^2\right]\Psi=0, \tag{15}$$

which is the governing equation of SH waves in a plate.[9] This observation indicates that when $\eta\approx 1$, guided SH waves in a hollow cylinder can be approximated as SH waves in a plate, which is practically much easier to handle.

III. NUMERICAL COMPUTATION AND DISCUSSION

In this section, numerical examples are presented for guided circumferential SH waves in carbon steel hollow cylinders of $\eta=0.2$, 0.5, and 0.8, respectively. The results of SH waves in a plate (corresponds to $\eta=1$) are also shown for comparison. In those calculations, the longitudinal and shear velocities of the carbon steel were chosen as $c_L=5900$ m/s, $c_T=3200$ m/s. For a given frequency f, numerical solutions of the nondimensional wave numbers $\hat{k}$ were obtained by applying the commonly used bisectional root-search method to the dispersion equation (11). The phase and group veloci-

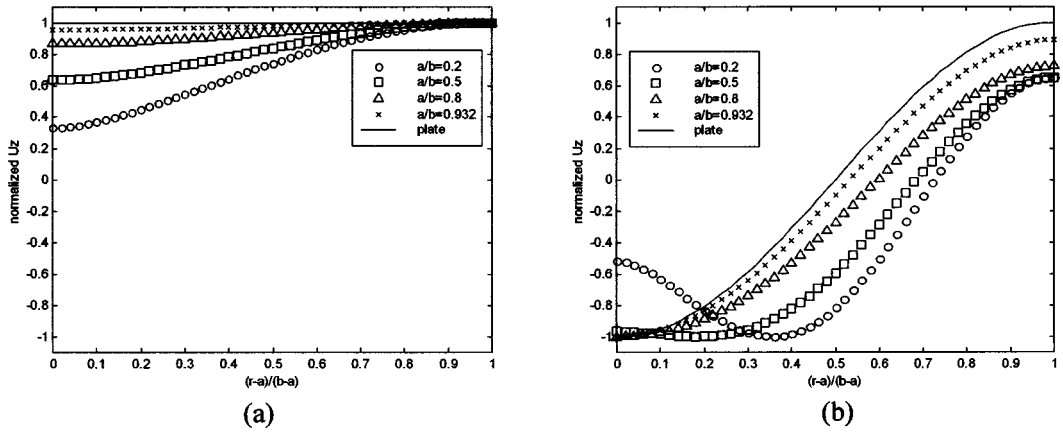


FIG. 4. Circumferential SH wave particle displacement distribution (a) n_0 mode at $fh=1$ and (b) n_1 mode at $fh=3$ in the wall of a hollow cylinder of $a/b=0.2$, 0.5, 0.8, and 0.932, respectively. SH waves in a plate are also shown for comparison.

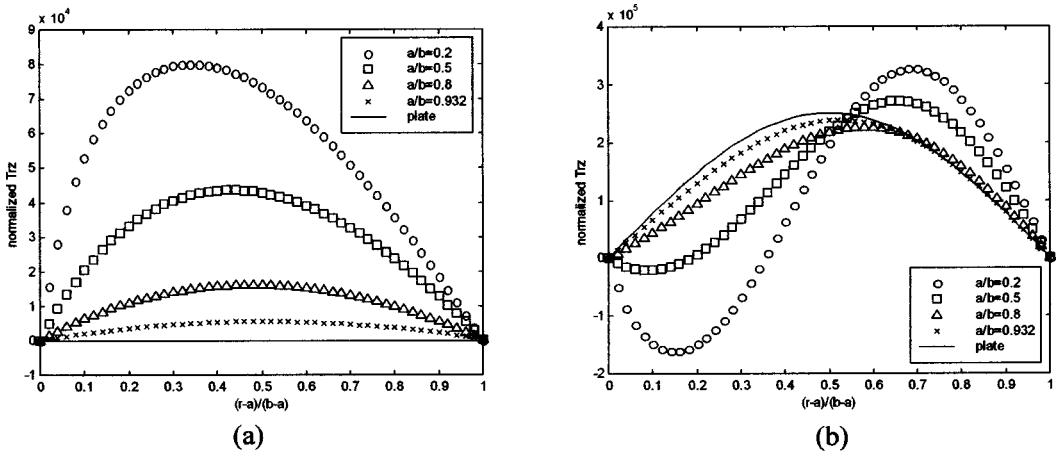


FIG. 5. Circumferential SH wave stress component τ_{rz} distribution (a) n_0 mode at $fh=1$ and (b) n_1 mode at $fh=3$ in the wall of a hollow cylinder of $a/b=0.2$, 0.5, 0.8, and 0.932, respectively. SH waves in a plate are also shown for comparison.

ties of the circumferential SH waves were obtained subsequently. Their dispersion curves are shown in Figs. 2(a) and (b), respectively. It is seen that the phase velocity dispersion curves of the circumferential SH waves resembles those in a plate, e.g., the lowest-order SH modes all start from a finite value, while the higher-order modes start from infinity with the cutoff frequency being the same as the corresponding SH mode in a plate. They all asymptotically approach the shear wave velocity of steel as fh tends to infinity. However, each SH mode in the hollow cylinder is shifted up and towards the right compared with that of a plate, and the lowest-order mode is no longer nondispersive as in the case of a plate. The group velocity dispersion curves of the hollow cylinder changed more dramatically with the change of η values. The lowest-order mode is no longer the fastest mode in energy propagation at high frequencies, and each high-order mode has a maximum value at a finite frequency. The smaller the η value, the higher the peak value and lower the fh product where that maximum occurs.

It is also seen from Figs. 2(a) and (b) that when the η value approaches 1, the circumferential guided SH wave dispersion curves approach those of a flat plate with the same wall thickness. To be quantitative, relative approximation errors $(c_{\text{pipe}}-c_{\text{plate}})/c_{\text{pipe}}$, where c denotes phase or group velocity, were calculated for each wave mode. For a 10-in. schedule-40 pipe (inner radius 127.25 mm, outer 136.53 mm, the corresponding $\eta\approx0.932$), the relative errors of using the plate model to calculate the phase and group velocities of the circumferential guided SH waves in the frequency range from 0.005 to 1 MHz are plotted in Figs. 3(a) and (b), respectively. It is seen that the relative errors are within 5% in the frequency region considered, although the error curves for the group velocity seem a bit rough due to the numerical derivation. Liu and Qu[7] did the same study for the circumferential guided plane–strain waves in a circular annulus versus Lamb waves in a plate. They drew very similar conclusions on using the plate model dispersion equation to approximate a cylindrical shell of large η value.

To further confirm the above observation and understand the propagation characteristics of the guided circumferential SH waves, the displacement field and stress field distribution in the hollow-cylinder wall of the first two modes are calculated from Eqs. (12), (13), and (14) for various η values. They are plotted in parts (a) and (b) of Figs. 4, 5, and 6, respectively. In these calculations, the fh value for the n_0 mode is arbitrarily chosen as 1 and the n_1 mode is 3; the practical case $\eta=0.932$ is also included in the plot. For better illustration, displacement field distributions are normalized

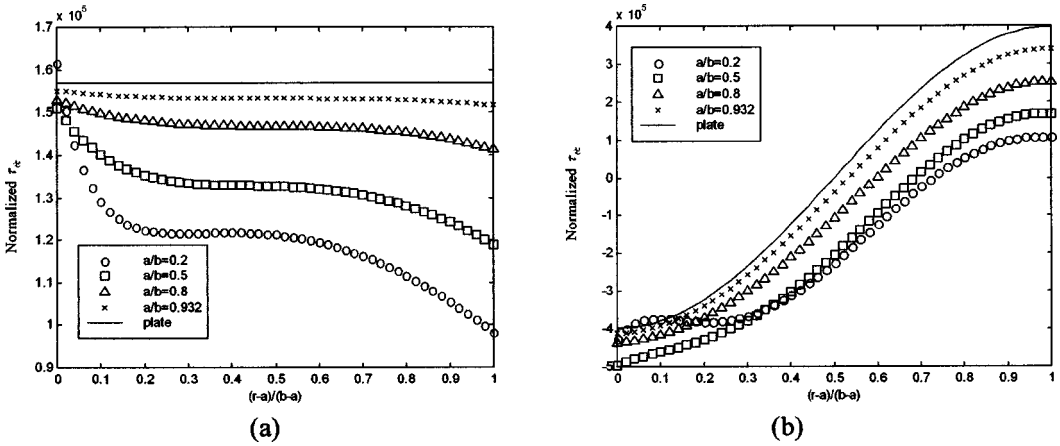


FIG. 6. Circumferential SH wave stress component $\tau_{\theta z}$ distribution (a) n_0 mode at $fh=1$ and (b) n_1 mode at $fh=3$ in the wall of a hollow cylinder of $a/b=0.2$, 0.5, 0.8, and 0.932, respectively. SH waves in a plate are also shown for comparison.

within ± 1 by each absolute maximum displacement value for each η; the stress components are then calculated corresponding to the normalized displacement. Figures 4(a) and (b) show the amplitude distribution of the particle displacements of the n_0 and n_1 mode guided circumferential SH waves, respectively. With $\eta \rightarrow 1$, the distribution was seen to approach monotonically that of a plate of the same fh value. The same trend can also be noticed in Figs. 5 and 6(a) and (b), which plot the stress components τ_{rz} and $\tau_{\theta z}$ of n_0 and n_1 modes, respectively. Once again, the plate model is a very accurate approximation to a hollow cylinder of $\eta \approx 1$.

On the other hand, when η becomes smaller, all the displacement and stress field distributions become more asymmetric with respect to the midplane of the hollow-cylinder wall. The lateral displacement of the n_0 mode circumferential SH wave has a larger value at the outer surface of the hollow cylinder [see Fig. 4(a)], while its stress component $\tau_{\theta z}$, which measures the interaction between particles in the wave propagation direction, is smaller [see Fig. 6(a)]. The interaction between particles in the radius direction τ_{rz} is no longer zero as is the case for a plate [see Fig. 5(a)]; it has maximum shifting toward the inner surface as $\eta \rightarrow 0$ and remains zero at the inner and outer free surfaces. For the n_1 mode, the zero crossings of both the displacement field u_z and the stress field $\tau_{\theta z}$ shift toward the outer surface as $\eta \rightarrow 0$ and more energy is concentrated near the inner surface [see Figs. 4(b) and 6(b)]; thus, the n_1 mode SH wave should have a better chance to detect defects located near the inner surface of a hollow cylinder.[10] The radius dependence of the stress component τ_{rz} also becomes complex [see Fig. 5(b)] as $\eta \rightarrow 0$, which is the result of a complex displacement profile across the wall thickness.

IV. CONCLUSIONS

Guided SH wave propagating in the circumferential direction of a hollow cylinder was studied. The dispersion equation as well as the displacement and stress field distribution across the wall of the hollow cylinder were derived analytically. Both the phase and group velocity dispersion curves and wave-field distribution are shown to be highly dependent on the wall-thickness-to-radius ratio of the hollow cylinder. When the ratio opts close to 1, they asymptotically approach that of a plate of the same frequency–thickness product. The relative error of phase and group velocities can serve as a quantitative measure of how well the plate approximation to the hollow cylinder is in wave analysis. When the ratio is small, precautions should be made on the mode and frequency selection in using the circumferential SH waves in defect detection due to the dispersion relation and wave-field distribution difference from that of a plate.

ACKNOWLEDGMENT

The authors acknowledge support from the Gas Technology Institute, Chicago IL.

[1] M. Hirao and H. Ogi, ''An SH-wave EMAT technique for gas pipeline inspection,'' NDT & E Int. **32**, 127–132 (1999).
[2] J. Gauthier, V. Mustafa, A. Chabbaz, and D. R. Hay, ''EMAT generation of horizontally polarized guided shear waves for ultrasonic pipe inspection,'' ASME Int. Pipeline Conf. **1**, 327–334 (1998).
[3] X. Zhao and J. L. Rose, ''Defect classification and sizing in a wave guide with Lamb and shear horizontal waves,'' Int. J. Solids Struct. **40**, 2645–2658 (2003).
[4] T. R. Meeker and A. H. Meitzler, ''Guided wave propagation in elongated cylinders and plates,'' in *Physical Acoustics, Principles and Methods* (Academic, New York, 1964), Vol. 1a.
[5] D. C. Gazis, ''Three-dimensional investigation of the propagation of waves in hollow circular cylinders. I. Analytical foundation,'' J. Acoust. Soc. Am. **31**, 568–573 (1959).
[6] D. C. Gazis, ''Three-dimensional investigation of the propagation of waves in hollow circular cylinders. II. Numerical results,'' J. Acoust. Soc. Am. **31**, 573–578 (1959).
[7] G. Liu and J. Qu, ''Guided circumferential waves in a circular annulus,'' J. Appl. Mech. **65**, 424–430 (1998).
[8] X. Zhao and J. L. Rose, ''Three-dimensional boundary element modeling for guided waves scattering from a defect,'' 30th Annual Conference on Rev. Prog. Quant. Nondestructive Eval., Green Bay, WI, 27 July–1 Aug., 2003.
[9] J. L. Rose, *Ultrasonic Waves in Solid Media* (Cambridge University Press, Cambridge, 1999).
[10] J. J. Ditri, J. L. Rose, and G. Chen, ''Mode selection criteria for defect optimization using lamb waves,'' Rev. Prog. Quant. Nondestr. Eval. **11**, 2109–2115 (1992).

Wave propagation and damping in linear viscoelastic laminates

P. J. Shorter[a)]
Vibro-Acoustic Sciences Inc, 12555 High Bluff Drive, Suite 250, San Diego, California 92130

(Received 16 May 2003; accepted for publication 1 February 2004)

This analysis is concerned with wave propagation and damping in linear viscoelastic laminates. A spectral finite-element method is developed and used to calculate the dispersion properties of the first few wave types of a given laminate; the proposed approach provides a robust and numerically efficient alternative to the transfer matrix method in certain applications. The proposed approach is also well suited to the calculation of the wave types of sections whose material properties vary continuously throughout the thickness of the section. A one-dimensional finite-element mesh is used to describe the through-thickness deformation of the laminate, and the dispersion equation for plane-wave propagation is formulated as a linear algebraic eigenvalue problem in wave number at each frequency of interest. The resulting eigenvectors and eigenvalues can be computed using standard numerical routines and used to investigate the dispersion characteristics of the propagating wave types of the section. The damping loss factor of each wave type is estimated from the cross-sectional strain energy distribution of the laminate. The proposed approach is well suited to modeling the structural-acoustic response of sandwich panels, constrained layer damping treatments, and general viscoelastic laminate sections in statistical energy analysis (SEA) codes. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1689342]

PACS numbers: 43.20.Bi, 43.55.Wk, 43.35.Mr, 43.20.Mv [RLW] Pages: 1917–1925

I. INTRODUCTION

Viscoelastic laminates have found application in many areas of structural acoustics ranging from constrained layer damping treatments to laminated windshields.[1–3] The widespread use of such laminates is due in part to the high levels of damping that can be attained when the cross-sectional properties of the laminate are chosen appropriately. A key requirement for determining the optimal cross-sectional properties of a given laminate is an accurate model of the structural-acoustic behavior of the laminate.

Typically, at low frequencies, a finite-element (FE) model provides a good description of the structural-acoustic behavior of the laminate. A simple displacement field is typically assumed through the thickness of the laminate, and low-order polynomials are used to describe the displacement field across each element.[4] At higher frequencies, the wavelengths of interest become small with respect to the lateral dimensions of the laminate (although, for structural-acoustic applications, the wavelengths are still typically large compared with the cross-sectional thickness of the laminate). The precision with which the lateral extent of a laminate (and its boundary conditions) are assumed to be known in an FE model then becomes unrealistic. In such instances, it becomes more appropriate to adopt a statistical analysis that accounts for uncertainty in the lateral dimensions of the laminate. The method of statistical energy analysis (SEA) provides such an approach.[5,6] In order to model a two-dimensional subsystem in (the wave approach to) SEA, it is necessary to determine the dispersion properties and damping loss factors of the propagating wave types of the subsystem.

Kerwin and Ungar provided an early analysis of the effect of a thin, constrained viscoelastic layer on the flexural waves of a panel.[7–10] The approach was based on an assumed low-order displacement field for the cross-sectional response of the laminate; the accuracy of such an approach is intrinsically linked to the displacement field assumed in the analysis. Several authors have described extensions to Kerwin and Ungar's analysis that use (low-order) displacement fields to describe the response of more general laminates (see Refs. 11 through 13, for example). Typically, the assumption of a low-order displacement field (or simple beam and plate theories for the individual layers) is required in such derivations in order to reduce analytical complexity. While such simplified analytical models can often provide physical insights into the behavior of certain laminates, the assumed displacement fields can often restrict the types of laminates that can be modeled. It is noted in passing that such methods are standard references in structural-acoustics textbooks that discuss the estimation of damping arising from viscoelastic materials.[3,14,15]

An alternative approach is to use an exact description of the through-thickness deformation of each layer of a laminate at a given frequency of interest. This approach is sometimes referred to as the transfer matrix method[16–18] (although the term global matrix method has also been used to describe a variant of this approach[18]). In such methods, the characteristic equation that describes free-wave propagation in a laminate takes the form of a nonlinear transcendental eigenvalue problem. Numerical root-finding algorithms are then usually employed to determine the dispersion properties of the free-wave types of the section.[18–20] Such algorithms typically search for sign changes or minima in the determinant of the global transfer matrix for various complex values of wave number.

While such an approach is (analytically) exact, from a

a)Electronic mail: pj.shorter@vasci.com

numerical point of view the approach often lacks robustness and careful attention needs to be paid to the step size and sampling in wave number space in order to ensure that no wave types have been missed. This is especially true for roots that are closely spaced or repeated (too coarse a sampling results in missed roots; too fine a sampling results in excessive computational expense). While such an approach can provide an accurate description of the response of a given laminate, the computational expense and robustness of the root-tracking algorithms employed can sometimes limit the usefulness of the approach for implementation in general-purpose structural-acoustic analysis codes (for which robustness and numerical efficiency are of utmost importance).

Consider, for example, a typical SEA model of the interior acoustic response of an automobile due to airborne excitation. Such a model may contain upwards of 60 or 70 different physical section properties, each representing different viscoelastic laminated glasses, panels with internal viscoelastic layers (metal–polymer–metal constructions), panels with baked-on mastics, and panels with constrained layer damping treatments. In a general-purpose SEA code, it is extremely important that the dispersion curves of such sections be automatically computed (and recomputed) with little or no user intervention and with a minimal amount of computational expense. Furthermore, in many instances, the material properties of a given viscoelastic layer are only available across a limited frequency range (or at various discrete frequencies). It is therefore highly desirable to have a method for computing the wave types at a fixed frequency, which does not require as a prerequisite, information about the wave types at neighboring frequencies (this constraint restricts the usefulness of root-searching algorithms which are based on frequency extrapolation routines[18]).

For most problems in structural acoustics, the cross-sectional (or through-thickness) displacement field of a viscoelastic laminate typically exhibits relatively few spatial oscillations across an individual layer of the laminate. It is therefore often possible to achieve a sufficiently precise description of the cross-sectional deformation of a laminate with a numerical method that employs an approximate description of the cross-sectional displacement field (which is, nevertheless, more precise than the displacement field assumed in simple plate or beam theories). Furthermore, for structural-acoustic applications, a precise description of the overall damping of a section (which precisely accounts for energy flows in the thickness direction of the laminate) is seldom necessary. For such applications, it is often acceptable to determine the overall section damping from consideration of the strain energy and loss factors of the section. Such observations lead one to the conclusion that, for structural-acoustic applications, a slight reduction in the numerical precision of the computed roots (through the use of an approximate numerical method), and an approximate inclusion of viscoelasticity, are entirely acceptable if improvements can be made to the robustness and computational efficiency of the overall method. In most practical applications, the lack of precision in one's knowledge of the properties of a viscoelastic material (and their temperature and frequency dependence) usually greatly exceeds any loss of precision due to the use of an approximate numerical method. It should be noted that such conclusions may not necessarily apply to applications in other areas of acoustics (such as ultrasonics and geophysics) and due consideration should be given to the trade-offs between accuracy, robustness, and numerical efficiency when considering the use of a given method.

The cross-sectional displacement field of the lower-order wave types of a viscoelastic laminate can be efficiently described in terms of a series of piecewise continuous basis functions (with shapes that are independent of frequency). While various choices of basis function exist, the use of low-order one-dimensional finite elements through the cross section provides a generic and flexible description of the displacement field that can be refined to achieve a desired level of accuracy. The previous conclusions have been reached by a number of authors in different contexts, and numerous studies therefore exist which employ a finite-element description of the cross section of a waveguide.

The use of a finite-element mesh to describe the cross-sectional deformation of a one-dimensional waveguide has been discussed extensively in studies concerned with wave propagation in linear elastic thin walled shells,[21] railway tracks,[22] rib stiffened plates,[23] rods,[24] and fluid filled pipes.[25] A recent overview of such approaches has been provided by Finnveden.[26] In each of the previous studies, the cross-sectional area of the waveguide is of finite extent and the finite elements are used to define a two-dimensional displacement field across the cross section of the waveguide.

The use of a one-dimensional finite-element mesh to describe the cross-sectional deformation of a laminate was first proposed by Dong and Nelson.[27,28] The analysis assumed a condition of plane strain and (neglecting shear waves) showed how the characteristic equation for free-wave propagation can be formulated as a linear eigenvalue problem in frequency. Further work demonstrated that the complex roots of the dispersion curves could be obtained by phrasing the characteristic equation as an eigenvalue problem in wave number.[29] A similar approach was later employed by Datta *et al.*[30] to investigate the dispersion curves of anisotropic laminates (although the authors referred to the resulting approach as a ''stiffness'' method rather than making explicit reference to the use of FE). Similar assumptions to those of Dong and Nelson were adopted regarding the state of plane strain in order to simplify the algebra. More recently, Xi *et al.*[31] have used a strip element method (essentially an FE-based method similar to the approach proposed by Dong and Nelson) to investigate the dispersion curves of elastic laminated composite cylinders. In this study, a full displacement field was employed for the cylinder and the characteristic equation formulated as a linear eigenvalue problem in frequency. While such an approach is useful for elastic laminates, the use of an eigenvalue problem in frequency is problematic for viscoelastic laminates due to the frequency dependence of the material properties of the viscoelastic layers.

In the present work, a one-dimensional finite-element mesh is used to describe the cross-sectional deformation of a linear viscoelastic laminate. The approach adopted here is

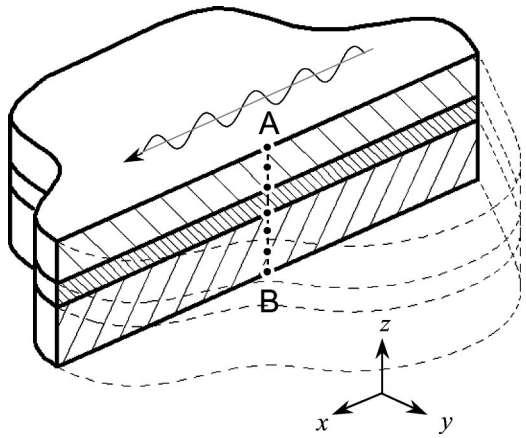


FIG. 1. A planar laminate with a one-dimensional finite-element mesh describing the displacement field between points A and B.

essentially the same as that proposed by Dong and Nelson; however, the current work considers a full three-dimensional displacement field within the laminate (rather than a state of plane strain). Linear elements are used to interpolate the cross-sectional displacement field and expressions are found for the kinetic and potential energy of the laminate at a given frequency of interest. The characteristic equation for free-wave propagation is formulated as a linear algebraic eigenvalue problem in wave number. The wave types for a given laminate at a fixed frequency can therefore be found numerically using standard algorithms for the solution of linear algebraic eigenproblems. The strain energy distribution through the section is then computed for each wave type and used to estimate the damping loss factor for each wave type.

II. CHARACTERISTIC EQUATION

A. Displacement field

Consider the viscoelastic laminate illustrated in Fig. 1. The laminate consists of an arbitrary number of solid linear viscoelastic and elastic layers. The current analysis assumes that the layers are planar and isotropic; however the extension to anisotropic and/or curved laminates is straightforward (as discussed in subsequent sections). It is assumed that a plane wave propagates in the positive x direction with frequency ω and wave number k so that the displacement field $\mathbf{d}$ at an arbitrary location within the laminate can be written as

$$\mathbf{d}(x,y,z,t)=[u(z)\quad v(z)\quad w(z)]^T\exp(i\omega t-ikx), \quad (1)$$

where u, v, and w define the displacements of the cross section in the x, y, and z directions, respectively. The continuous cross-sectional displacement field can be approximated with various finite elements so that

$$\begin{bmatrix} u(z) \\ v(z) \\ w(z) \end{bmatrix} = \begin{bmatrix} \mathbf{N}_i(z) & 0 & 0 \\ 0 & \mathbf{N}_i(z) & 0 \\ 0 & 0 & \mathbf{N}_i(z) \end{bmatrix} \begin{bmatrix} \mathbf{q}_u \\ \mathbf{q}_v \\ \mathbf{q}_w \end{bmatrix} = \mathbf{N}(z)\mathbf{q}_0, \quad (2)$$

where the matrix $\mathbf{N}$ contains the element shape functions and the vector $\mathbf{q}_0$ contains the complex amplitudes of the nodal degrees of freedom. The displacement field can therefore be written as

$$\mathbf{d}(x,y,z,t)=\mathbf{N}(z)\mathbf{q}_0\exp(i\omega t-ikx). \quad (3)$$

B. Kinetic energy

The time-average kinetic energy T in a volume Ω of the laminate is given by

$$T=\frac{\omega^2}{4}\int_\Omega \rho(\mathbf{x})\mathbf{d}^H(\mathbf{x})\mathbf{d}(\mathbf{x})d\mathbf{x}, \quad (4)$$

where $.^H$ denotes the complex conjugate or Hermitian transpose and ρ is the local mass density. Substituting for the discretized displacement field gives

$$T=\frac{\omega^2}{4}\mathbf{q}_0^H\mathbf{M}\mathbf{q}_0\,dx\,dy, \quad (5)$$

where the mass matrix $\mathbf{M}$ is defined by

$$\mathbf{M}=\int_z \rho(z)\mathbf{N}^H(z)\mathbf{N}(z)dz. \quad (6)$$

The mass matrix may be assembled from various elemental contributions as discussed in the Appendix.

C. Potential energy

The (engineering) strain at a particular point in the laminate is related to the displacement field by[32]

$$\mathbf{e}(\mathbf{x},t)=\begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \gamma_{xy} \\ \gamma_{xz} \\ \gamma_{yz} \end{bmatrix} = \begin{bmatrix} \partial/\partial x & 0 & 0 \\ 0 & \partial/\partial y & 0 \\ 0 & 0 & \partial/\partial z \\ \partial/\partial y & \partial/\partial x & 0 \\ \partial/\partial z & 0 & \partial/\partial x \\ 0 & \partial/\partial z & \partial/\partial y \end{bmatrix} \mathbf{d}. \quad (7)$$

Substituting for the discretized displacement field gives

$$\mathbf{e}(\mathbf{x},t)=\mathbf{B}(k,z)\mathbf{q}_0\exp(i\omega t-ikx), \quad (8)$$

where the strain–displacement matrix $\mathbf{B}(k,z)$ is given by

$$\mathbf{B}(k,z)=\begin{bmatrix} -ik\mathbf{N}(z) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \dfrac{\partial\mathbf{N}(z)}{\partial z} \\ 0 & -ik\mathbf{N}(z) & 0 \\ \dfrac{\partial\mathbf{N}(z)}{\partial z} & 0 & -ik\mathbf{N}(z) \\ 0 & \dfrac{\partial\mathbf{N}(z)}{\partial z} & 0 \end{bmatrix}. \quad (9)$$

The stress at a given point is related to the strain by the constitutive law of the material so that

$$\mathbf{s}=\mathbf{D}(z)\mathbf{e}, \quad (10)$$

For an isotropic material the matrix $\mathbf{D}(z)$ is given by[32]

$$\mathbf{D}(z)=\begin{bmatrix} a & b & b & 0 & 0 & 0 \\ b & a & b & 0 & 0 & 0 \\ b & b & a & 0 & 0 & 0 \\ 0 & 0 & 0 & c & 0 & 0 \\ 0 & 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 0 & c \end{bmatrix}, \tag{11}$$

where the material constants in the previous expression are defined by

$$a=\frac{E(1-\nu)}{(1+\nu)(1-2\nu)}; \quad b=\frac{E\nu}{(1+\nu)(1-2\nu)};$$

$$c=\frac{E}{2(1+\nu)}. \tag{12}$$

For a linear viscoelastic material the material constants are frequency dependent and complex.[33] The time-average strain energy U in a volume Ω of the laminate is given by

$$U=\frac{1}{4}\mathbf{q}_0^H\mathbf{K}(k)\mathbf{q}_0\,dx\,dy, \tag{13}$$

where the stiffness matrix $\mathbf{K}$ is defined by

$$\mathbf{K}(k)=\int_z \mathbf{B}^H(k,z)\mathbf{D}(z)\mathbf{B}(k,z)dz. \tag{14}$$

Since the strain-displacement matrix only contains terms linear in k, the stiffness matrix in Eq. (14) can be written in the form

$$\mathbf{K}(k)=\mathbf{K}_2k^2+\mathbf{K}_1k+\mathbf{K}_0, \tag{15}$$

where the matrices $\mathbf{K}_i$ are independent of wave number. The stiffness matrix may be assembled from various elemental contributions as discussed in the Appendix. For a dissipative material the strain energy defined by Eq. (13) consists of a real component and an imaginary component. The real component describes the time-average strain energy throughout the section, while the imaginary component is related to the time-average power dissipated by the section.

D. Lagrange's equations

Equations of motion for the cross section can be formulated by inserting the kinetic and potential energies defined in Eqs. (5) and (13) into the time-average form of Lagrange's equations.[34] In general, the nonconservative form of Lagrange's equations should be used to account for dissipation through the cross section. However, the following analysis adopts a simplified approach in which: (i) the equations of motion are generated assuming that the laminate is conservative and (ii) the resulting cross-sectional strain energy distribution is used to estimate the power dissipated by the section. This assumption is consistent with the requirements of the intended application and has the advantage that it simplifies the classification of the roots into propagating and evanescent wave types (the subsystems in an SEA model are usually associated with the propagating wave types of a given section). The assumption is likely to be valid if the cross-sectional strain energy distribution of a given propagating wave is not significantly modified by increasing levels of damping. It should be noted that this assumption is implicit in previous studies which assume a fixed displacement field for the cross section that is independent of damping.

Inserting Eq. (5) and the real part of Eq. (13) into the conservative form of Lagrange's equations results in the following characteristic equation:

$$[\mathbf{K}_r(k)-\omega^2\mathbf{M}]\mathbf{q}_0=0. \tag{16}$$

The subscript r in the previous expression is used to indicate that the real part of the constitutive matrix $\mathbf{D}$ is used when generating the stiffness matrix.

If a particular wave number k is specified, then Eq. (16) takes the form of a generalized eigenvalue problem for ω. If a particular frequency ω is specified, then Eq. (16) takes the form of a generalized quadratic eigenvalue problem for k. The frequency dependence of the material properties of a viscoelastic laminate means that the latter must be used in the current analysis. The quadratic eigenvalue problem can be transformed into a linear eigenvalue problem by employing the transformations[23]

$$\mathbf{q}_2=k\mathbf{q}_1; \quad \mathbf{q}_1=k\mathbf{q}_0. \tag{17}$$

Inserting the previous expression into Eq. (16) results in the generalized eigenvalue problem given by

$$\begin{bmatrix} 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \\ \mathbf{C}_0 & \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix}\begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix}-k\begin{bmatrix} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix}=0, \tag{18}$$

where

$$\mathbf{C}_2=\mathbf{K}_{r,2}; \quad \mathbf{C}_1=\mathbf{K}_{r,1}; \quad \mathbf{C}_0=\mathbf{K}_{r,0}-\omega^2\mathbf{M}. \tag{19}$$

Equation (18) can be converted to a standard form by eliminating $\mathbf{q}_2$ from the previous equations to give

$$\begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{C}_2^{-1}\mathbf{C}_0 & -\mathbf{C}_2^{-1}\mathbf{C}_1 \end{bmatrix}\begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \end{bmatrix}-k\begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \end{bmatrix}=0. \tag{20}$$

The matrix $\mathbf{C}_2$ can be inverted if the material constants a and c defined in Eq. (12) are nonzero (this will be the case for materials with a nonzero Young's modulus and a Poisson's ratio in the range $0<\nu<0.5$). Equation (20) phrases the characteristic equation as a standard linear algebraic eigenvalue problem and is the main result of this section.

The eigenvalues and eigenvectors of Eq. (20) may be obtained using standard routines. The cross-sectional displacement field of the lower-order wave types of a given laminate can usually be described using a relatively small number of elements; the solution of the previous eigenproblem is therefore computationally efficient. The eigenvalues can be separated into: (i) those that are purely real (representing propagating waves); (ii) those that are purely imaginary (representing nonoscillating evanescent waves); and (iii) those that are complex (representing oscillating evanescent waves).[22] The sign of the real and imaginary parts of the eigenvalues indicates the direction of travel of a given wave. The eigenvectors indicate the cross-sectional mode shape associated with a given wave (at a particular frequency of in-

terest). In general, the eigenvectors are complex (due to variations in the relative phase of the displacement field throughout the cross section).

E. Damping loss factor

The damping loss factor associated with a given propagating wave is defined by[3]

$$\eta = \frac{P_{\text{diss}}}{\omega(T+U)}, \tag{21}$$

where P_{diss} is the total time-average power dissipated by the wave. Assuming that the loss factor of the nth layer of the laminate is uniform through the thickness of the layer, the total dissipated power can be related to the strain energy of the individual layers by

$$P_{\text{diss}} = \sum_n 2\omega\eta_n U_n, \tag{22}$$

where the summation n is over the number of layers in the laminate. Substituting Eqs. (5), (13), and (22) into (21) gives

$$\eta = \frac{\Sigma_n 2\eta_n \mathbf{q}_0^H \mathbf{K}_r^{(n)} \mathbf{q}_0}{\mathbf{q}_0^H (\mathbf{K}_r + \omega^2 \mathbf{M}) \mathbf{q}_0}, \tag{23}$$

where $\mathbf{K}^{(n)}$ is the contribution to the assembled stiffness matrix arising from the nth layer of the laminate. The damping loss factor for the kth propagating wave can then be written as

$$\eta_k = \frac{\Sigma_n \eta_n \mathbf{P}_k^H \mathbf{K}_r^{(n)} \mathbf{P}_k}{\mathbf{P}_k^H \mathbf{K}_r \mathbf{P}_k}, \tag{24}$$

where $\mathbf{P}_k$ is the eigenvector associated with the kth propagating wave. Equation (24) provides an estimate of the damping loss factor for a given propagating wave type and is the main result of this section. The expression is in agreement with the widely known result that the damping loss factor of a laminate is maximized by maximizing the proportion of strain energy contained within layers with high damping loss factors.[3]

III. NUMERICAL EXAMPLES

The use of the previous expressions is demonstrated in the following sections with a number of simple numerical examples.

A. Thin plate

A 1-mm-thick isotopic steel plate was modeled with material properties given by: $E = 210\text{e}9$ Pa, $\nu = 0.3125$, $\rho = 7800$ kg/m^3. Six linear elements were used to describe the cross-sectional displacement of the plate (resulting in 21 nodal degrees of freedom). The wave numbers of the three propagating wave types of the section were computed from Eq. (20) and are plotted in Fig. 2. The analytical estimates of the wave numbers of the propagating waves of a thin isotropic plate are given by[35]

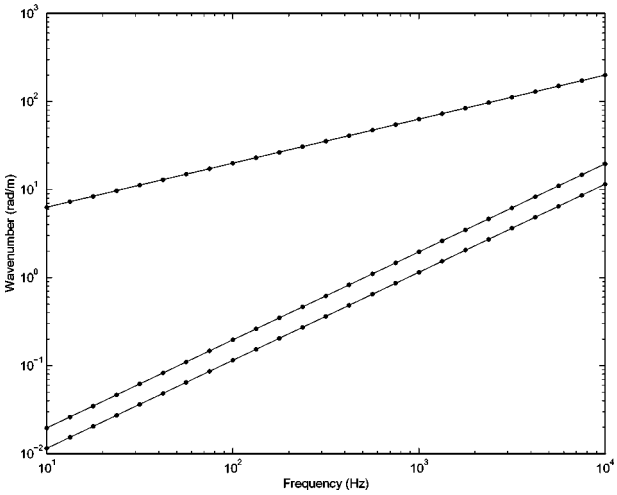


FIG. 2. Dispersion curves for propagating wavetypes of a 1-mm-thick steel panel: solid line, analytical result; dotted line, FE computed result.

$$k_f^4 = \frac{12(1-\nu^2)\rho}{Eh^2}\omega^2; \quad k_s^2 = \frac{2\rho(1+\nu)}{E}\omega^2;$$

$$k_e^2 = \frac{\rho(1-\nu^2)}{E}\omega^2. \tag{25}$$

The analytical wave numbers are also plotted in Fig. 2 and are in good agreement with the FE-computed wave numbers. A convergence study was performed which indicated that, for this example, the FE results converge using one or two elements per layer. However, since the computational expense of the approach is not significant, the following examples employ six linear elements per layer.

B. Constrained layer damping treatment

Kerwin[7] has presented measurements and predictions of the flexural damping loss factor of a thin aluminum beam with an attached constrained layer damping treatment. Curves are given for the frequency dependence of the real and imaginary parts of the shear modulus of the viscoelastic material at various temperatures (Fig. 6 of Ref. 7). The damping loss factors measured by Kerwin for three different constraining layer thicknesses have been plotted in Fig. 3 (compare with Fig. 7 of Ref. 7). The flexural damping loss factors for the three configurations were predicted using Kerwin's Eq. IV-5 (using the material property data obtained at a temperature of 25 °C) and plotted in Fig. 3. The flexural damping loss factors were also computed using the FE formulation described in the current work (using typical material property values for aluminum). The FE-computed flexural damping loss factors have been plotted in Fig. 3 and are in good agreement with the loss factors predicted by Kerwin (especially at lower frequencies where the assumption of a thin constraining layer used in the derivation of Kerwin's Eq. IV-5 are likely to be most valid). It can be seen that discrepancies still exist with Kerwin's measured damping loss factors above 1 kHz (these discrepancies may, however, be re-

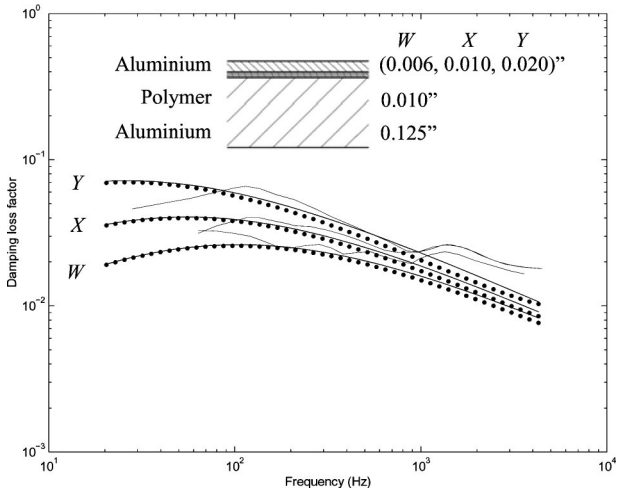


FIG. 3. Flexural damping loss factor of a panel with a constrained layer damping treatment. Results plotted for three different constraining layer thicknesses (from Fig. 7 of Ref. 7): solid black line, damping loss factor calculated from Eq. IV-5 of Ref. 7; solid gray line, experimental measurements from Ref. 7; dotted line, FE-computed results.

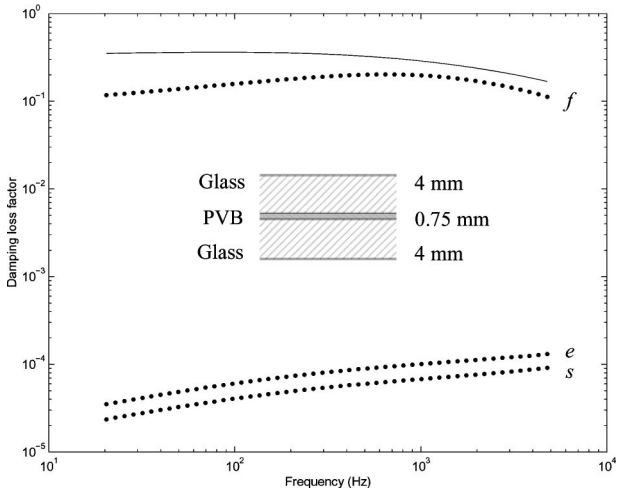


FIG. 5. Damping loss factors for typical automotive glass laminate: solid line, damping loss factor calculated from Eq. IV-5 of Ref. 7; dotted lines, FE-computed damping loss factor for flexural, extensional, and shear waves.

lated to the use of extrapolated material property data and the differences in the wave types of a plate and a finite-width beam).

C. Glass laminate

Kerwin's original analysis was intended for thin constrained layer damping treatments and the assumptions that were made in the derivation of Eq. IV-5 are clearly stated in Ref. 7. However, it is of interest to explore the errors that one encounters when applying this equation to viscoelastic laminates with thicker constraining layers. A typical automotive glass laminate was chosen which consists of two 4-mm-thick sheets of glass separated by a 0.75-mm-thick layer of polyvinyl butyaryl (PVB). The approximate frequency dependence of the viscoelastic material properties of PVB is shown in Fig. 4. The elastic material properties of the glass are taken to be: $E=62\text{e}9$ Pa, $\nu=0.24$, $\rho=2300$ kg/m^3. To emphasize the dissipative effects of the viscoelastic interlayer the glass is assumed to have an internal loss factor of zero.

The FE-computed loss factors for flexural, shear, and extensional waves are plotted in Fig. 5 and compared with the flexural loss factor predicted using Eq. IV-5 of Ref. 7. It can be seen that, for this example, the assumptions made in Ref. 7 result in an overprediction of the flexural loss factor by a factor of approximately 3 at low frequencies. It can also be seen that the viscoelastic interlayer does not provide a significant amount of damping to extensional and shear waves. The strain energy for these waves is concentrated in the glass layers, and only a relatively small amount of strain energy is stored in the interlayer (in practice, the loss factors of the shear and extensional waves for such a section will be governed by the inherent material loss factor of the glass layers). The current FE approach, can, however be used to investigate the additional damping that results from the use of asymmetric sections with an arbitrary number of layers.

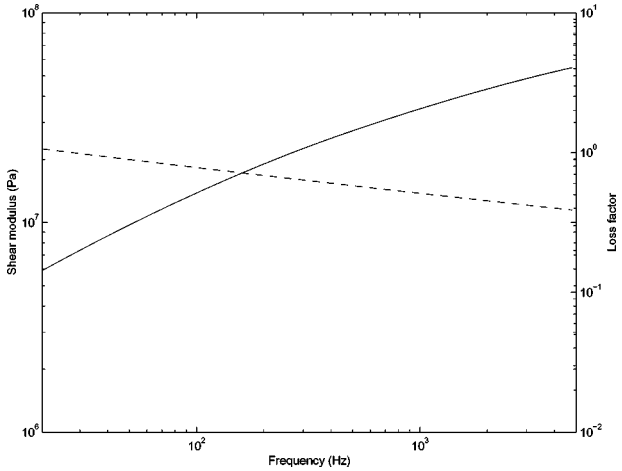


FIG. 4. Approximate material property data for polyvinyl butyaryl (PVB) as a function of frequency: solid line, shear modulus; dashed line, loss factor (ν and ρ assumed to be frequency independent and equal to 0.49 and 1000 kg/m^3, respectively).

D. Sandwich panel

As a final example, the FE approach is used to calculate the wave types of a typical elastic sandwich panel. Most sandwich panel cores are orthotropic; however, for simplicity the following example considers an isotropic core. The effective isotropic properties of the core used in this example are given by: $E=30\text{e}6$ Pa, $\nu=0.2$, $\rho=48$ kg/m^3. The skins are taken to be aluminum with isotropic properties given by: $E=71\text{e}9$ Pa, $\nu=0.3296$, $\rho=2700$ kg/m^3. A symmetric sandwich construction is considered with 0.6-mm-thick skins and a 15-mm-thick core.

The dispersion curves of the propagating wave types of the sandwich panel were calculated using the FE approach and are plotted in Fig. 6. The cross-sectional mode shapes of the various propagating waves are illustrated in Fig. 7. Below approximately 5 kHz, three propagating waves exist corresponding to flexural (a), shear (c), and extensional (e) waves. Around 5 kHz two additional propagating waves cut on which involve an out-of-phase motion of the skins of the

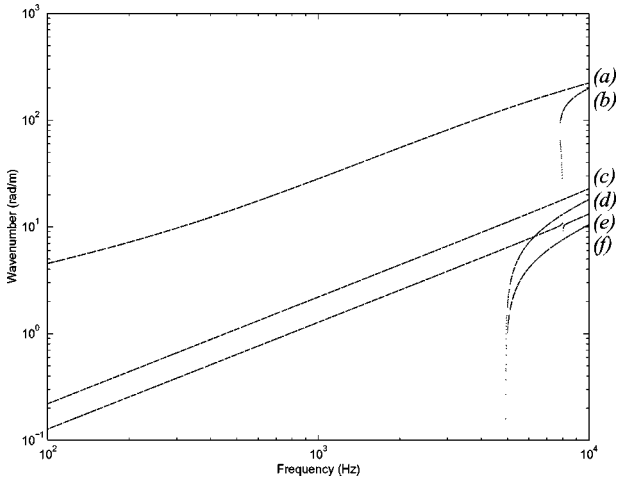


FIG. 6. Dispersion curves for wave types of a typical sandwich panel (cross-sectional mode shapes shown in Fig. 7).

panel in the in-plane direction (d) and (f). Around 8 kHz two additional propagating waves cut on. These waves cut on with a finite wave number; however, one wave has a negative group velocity and only exists over a very narrow frequency range; the other wave has a positive group velocity and exists over a much broader frequency range (b). Both waves involve an out-of-phase flexural motion of the skins of the panel.

The detailed behavior of the dispersion curves around 8 kHz predicted using the FE approach is plotted in Fig. 8(a). The complexity of the dispersion curves (and the veering of the extensional and flexural wave types) is clearly evident. For many applications, the structural-acoustic performance of a given section is adequately described by consideration of the first three wave types of the section. However, analysis of the higher-order wave types of a laminate is often important for structural-acoustic applications and can lead to novel laminate designs with improved structural-acoustic performance (as demonstrated by Moore and Lyon[36]).

To provide a comparison with the FE-computed results, a transfer matrix approach was also used to investigate the dispersion curves over the same region of wave number space. The transfer matrix approach that was used in this example[17] is primarily intended for modeling the response of poroelastic noise control treatments to acoustic excitation and therefore neglects shear wave behavior within each layer (by assuming a two-dimensional displacement field within each layer). The sign of the real part of the determinant of the global transfer matrix is plotted in Fig. 8(b) over a grid of 40 000 points in wave number space. It can be seen that the sign changes of the determinant of the transfer matrix reveal the same dispersion behavior predicted by the FE-based approach (albeit without shear waves for the reasons discussed previously).

This example also serves to highlight the problems one encounters when searching for the roots of the characteristic equation numerically using the transfer matrix approach. Consider, for example, the problem of determining the free-wave types of the sandwich section at a single discrete frequency of 8 kHz. At 8 kHz, the first three propagating wave

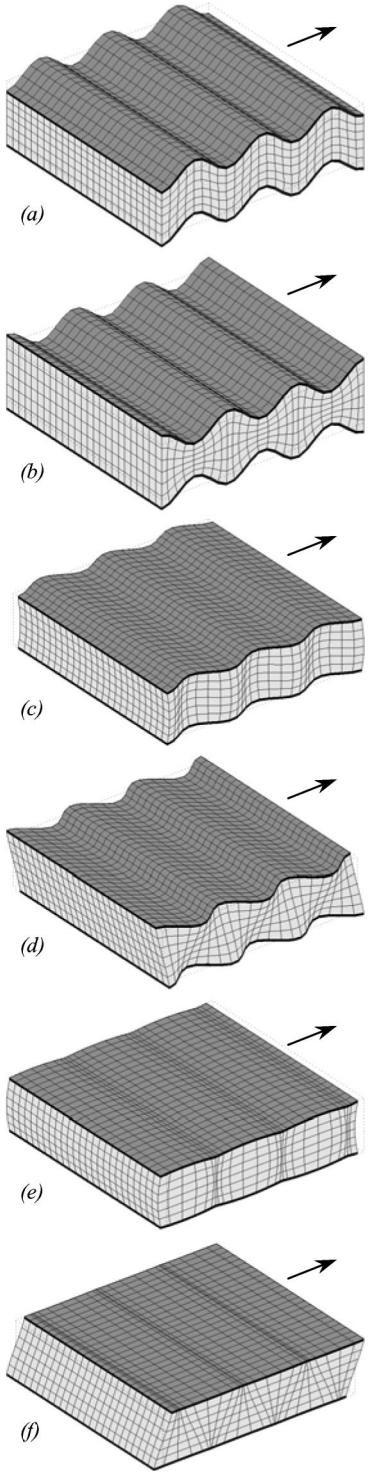


FIG. 7. Propagating wave types in a typical sandwich panel at 9 kHz calculated using 18 one-dimensional finite elements (plots normalized to have equal wave number).

types have wave numbers that are within approximately 2 rad/m of each other, while the flexural wave has a wave number of approximately 190 rad/m. Without knowledge of the roots at adjacent frequencies (and in the absence of algorithms which employ heuristic search strategies), an initial sample spacing of less than 1 rad/m would be required in order to avoid missing the second and third roots. Even when employing such an approach, there would be no guarantee

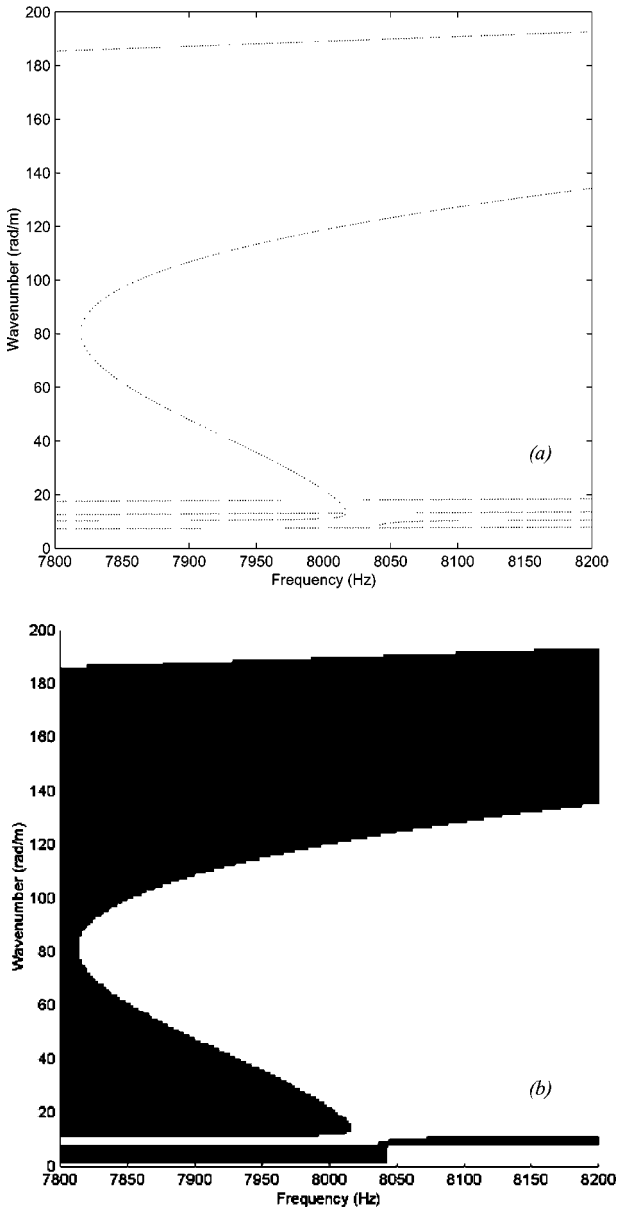


FIG. 8. Detailed view of dispersion curve in Fig. 6 at cut on of asymmetric bending wave: (a) computed using FE approach; (b) sign of the real part of the determinant of the characteristic equation computed using a transfer matrix approach over a grid of 40 000 points in wave number space.

that all the required roots had been extracted. For example, a root-searching strategy that simply extracted the first five propagating wave types (based on the observed dispersion behavior between 5 and 7 kHz) would fail to extract the correct number of propagating wave types between 7820 and 8020 Hz (where seven propagating wave types exist over a narrow frequency range). The problem is exacerbated if the section is viscoelastic; an exact transfer matrix approach then requires that the root search be performed over both real and imaginary values of wave number (for each discrete frequency of interest). If one adopts a simple sampling approach to ensure that no roots are missed (and assumes similar ranges for the imaginary and real values of wave number), then the initial grid search would require at least 35 000 sample points. For general-purpose structural-acoustics codes, the advantages of a numerical method in which all the roots (and associated damping loss factors) are obtained in a robust and computationally efficient manner by solving a standard algebraic eigenvalue problem, become readily apparent.

IV. CONCLUDING REMARKS

This analysis has discussed an efficient numerical method for computing the lower-order wave types and damping loss factors of a viscoelastic laminate. A one-dimensional finite element mesh was used to describe the cross-sectional displacement field of the laminate and the characteristic equation for free-wave propagation was formulated as a linear algebraic eigenvalue problem in wave number [Eq. (20)]. The damping loss factors of the section were estimated by computing the strain energy distribution throughout the section for various propagating wave types of interest [Eq. (24)]. Validation examples showed the method to be in good agreement with previous work for simple laminates. The main contribution of the current work is the extension of existing methods to include the full three-dimensional displacement field and the consideration of viscoelasticity.

Mass and stiffness matrices were presented for planar isotropic layers. The analysis can, however, be extended to encompass anisotropic curvilinear panels through the use of the appropriate strain–displacement and stress–strain matrices in Eqs. (7) and (11) as a function of heading. The current analysis can also be used to examine wave propagation in sections in which the material properties vary continuously throughout the section thickness. The main assumption employed in the current analysis is that a conservative form of Lagrange's equations can be used to develop equations of motion for the cross section. The assumption may be relaxed through the use of the nonconservative form of Lagrange's equations. While such an extension is unlikely to be required for most structural-acoustic applications, it may enable the current approach to be applied to certain wave propagation problems encountered in ultrasonics and geophysics applications.

ACKNOWLEDGMENTS

This work was funded by Vibro-Acoustic Sciences Inc. Technical discussions with Bryce Gardner, Paul Bremner, Vincent Cotoni, and Robin Langley are gratefully acknowledged.

APPENDIX: MASS AND STIFFNESS MATRICES

The isoparametric mass and stiffness matrices for a linear isotropic two-noded element of length L with degrees of freedom $[u_a\ u_b\ v_a\ v_b\ w_a\ w_b]^T$ are Hermitian and given by

$$\mathbf{M}=\begin{bmatrix}\mathbf{M}_i & 0 & 0\\ 0 & \mathbf{M}_i & 0\\ 0 & 0 & \mathbf{M}_i\end{bmatrix};\quad \mathbf{M}_i=\rho L\begin{bmatrix}1/3 & 1/6\\ 1/6 & 1/3\end{bmatrix}. \tag{A1}$$

$$\mathbf{K}_{r,2}=\frac{L}{6}\begin{bmatrix} 2a & a & 0 & 0 & 0 & 0 \\ & 2a & 0 & 0 & 0 & 0 \\ & & 2c & c & 0 & 0 \\ & & & 2c & 0 & 0 \\ & \textit{herm.} & & & 2c & c \\ & & & & & 2c \end{bmatrix}; \tag{A2}$$

$$\mathbf{K}_{r,1}=\begin{bmatrix} 0 & 0 & 0 & 0 & -\frac{i}{2}(b-c) & \frac{i}{2}(b+c) \\ & 0 & 0 & 0 & -\frac{i}{2}(b+c) & \frac{i}{2}(b-c) \\ & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & \textit{herm.} & & & 0 & 0 \\ & & & & & 0 \end{bmatrix}; \tag{A3}$$

$$\mathbf{K}_{r,0}=\frac{1}{L}\begin{bmatrix} c & -c & 0 & 0 & 0 & 0 \\ & c & 0 & 0 & 0 & 0 \\ & & c & -c & 0 & 0 \\ & & & c & 0 & 0 \\ & \textit{herm.} & & & a & -a \\ & & & & & a \end{bmatrix}. \tag{A4}$$

[1] E. Kerwin and E. Ungar, ''Requirements imposed on polymeric materials by structural damping applications,'' in *Sound and Vibration Damping with Polymers*, edited by R. Corsaro and L. Sperling (American Chemical Society, Washington, D.C., 1990), Chap.17, pp. 317–345.

[2] M. Rao, ''Recent applications of viscoelastic damping for noise control in automobiles and commercial airplanes,'' J. Sound Vib. **262**, 457–474 (2003).

[3] E. E. Ungar, ''Structural damping'' in *Noise and Vibration Control Engineering*, edited by L. Beranek and I. Ver (Wiley, New York, 1992), Chap.12, pp. 451–481.

[4] T. S. Pagianakos and D. A. Saravanos, ''Mechanics and finite elements for the damped characteristics of curvilinear laminates and composite shell structures,'' J. Sound Vib. **263**, 399–414 (2003).

[5] R. Lyon, *Statistical Energy Analysis of Dynamical Systems* (MIT Press, Cambridge, MA, 1975).

[6] F. Fahy, ''Statistical energy analysis: A critical overview,'' Proc. R. Soc. London, Ser. A **346**, 429–554 (1994).

[7] E. M. Kerwin, ''Damping of flexural waves by a constrained viscoelastic layer,'' J. Acoust. Soc. Am. **31**(7), 952–962 (1959).

[8] E. E. Ungar, ''Loss factors of viscoelastically damped beam structures,'' J. Acoust. Soc. Am. **34**(8), 1082–1089 (1962).

[9] E. E. Ungar and E. M. Kerwin, ''Loss factors of viscoelastic systems in terms of energy concepts,'' J. Acoust. Soc. Am. **34**(7), 954–957 (1962).

[10] E. E. Ungar and E. M. Kerwin, ''Plate damping due to thickness deformations in attached viscoelastic layers,'' J. Acoust. Soc. Am. **36**(2), 386–392 (1964).

[11] D. J. Mead and S. Markus, ''The forced vibration of a three-layer, damped sandwich beam with arbitrary boundary conditions,'' J. Sound Vib. **10**(2), 163–175 (1969).

[12] D. J. Mead, ''A comparison of some equations for the flexural vibration of damped sandwich beams,'' J. Sound Vib. **83**(3), 363–377 (1982).

[13] K. Heron, ''Predictive SEA and anisotropic panels,'' Proceedings of ISMA, Leuven, Belgium (2002).

[14] M. Crocker, *Handbook of Acoustics* (Wiley, New York, 1998).

[15] D. Jones, *Handbook of Viscoelastic Vibration Damping* (Wiley, New York, 2001).

[16] Y. K. Lin, *Probabilistic Theory of Structural Dynamics* (McGraw-Hill, New York, 1967).

[17] J. F. Allard, *Propagation of Sound in Porous Media* (Elsevier, Amsterdam, 1993).

[18] M. Lowe, ''Matrix techniques for modeling ultrasonic waves in multilayered media,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **42**(4), 525–542 (1995).

[19] G. Coquin, ''Attenuation of guided waves in isotropic viscoelastic materials,'' J. Acoust. Soc. Am. **36**(6), 1074–1080 (1964).

[20] M. Castaings and B. Hosten, ''Guided waves propagating in sandwich structures made of anisotropic, viscoelastic, composite materials,'' J. Acoust. Soc. Am. **113**(5), 2622–2634 (2003).

[21] L. Gavric, ''Finite element computation of dispersion properties of thin-walled waveguides,'' J. Sound Vib. **173**(1), 113–124 (1994).

[22] L. Gavric, ''Computation of propagative waves in free rail using a finite element technique,'' J. Sound Vib. **185**(3), 531–543 (1995).

[23] U. Orrenius and S. Finnveden, ''Calculation of wave propagation in rib-stiffened plate structures,'' J. Sound Vib. **198**(2), 203–224 (1996).

[24] T. Mazuch, ''Wave dispersion modelling anisotropic shells and rods by the finite element method,'' J. Sound Vib. **198**(4), 429–438 (1996).

[25] S. Finnveden, ''Spectral finite element analysis of the vibration of straight fluid-filled pipes with flanges,'' J. Sound Vib. **199**(1), 125–154 (1997).

[26] S. Finnveden, ''Finite element techniques for the evaluation of energy flow parameters,'' Proceedings of NOVEM, Lyon (2000).

[27] S. Dong and R. Nelson, ''On natural vibrations and waves in laminated orthotropic plates,'' J. Appl. Mech. **39**, 739–745 (1972).

[28] R. Nelson and S. Dong, ''High frequency vibrations and waves in laminated orthotropic plates,'' J. Sound Vib. **30**(1), 33–44 (1973).

[29] S. Dong and K. Huang, ''Edge vibrations in laminated composite plates,'' J. Appl. Mech. **52**, 433–438 (1985).

[30] S. Datta, A. Shah, R. Bratton, and T. Chakraborty, ''Wave propagation in laminated composite plates,'' J. Acoust. Soc. Am. **83**(6), 2020–2026 (1988).

[31] Z. Xi, G. Liu, K. Lam, and H. Shang, ''Dispersion and characteristic surfaces of waves in laminated composite circular cylindrical shells,'' J. Acoust. Soc. Am. **108**(5), 2179–2186 (2000).

[32] R. D. Cook, D. Malkus, and M. Plesha, *Concepts and Applications of Finite Element Analysis* (Wiley, New York, 1989).

[33] P. Grootenhuis, ''The control of vibrations with viscoelastic materials,'' J. Sound Vib. **11**(4), 421–433 (1970).

[34] P. Morse and K. Ingard, *Theoretical Acoustics* (McGraw-Hill, New York, 1986), p. 251

[35] L. Cremer and M. Heckl, *Structure-Borne Sound* (Springer, Berlin, 1988).

[36] J. Moore and R. Lyon, ''Sound transmission loss characteristics of sandwich panel constructions,'' J. Acoust. Soc. Am. **89**(2), 777–791 (1991).

Acoustic scattering from rigid bodies of arbitrary shape—Double layer formulation

B. Chandrasekhar and Sadasiva M. Rao[a)]
Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore—560 012, India

(Received 1 July 2003; revised 7 February 2004; accepted 17 February 2004)

In this work, a numerical method is presented, based on the well-known method of moments (MoM), to calculate the acoustic fields scattered by a three-dimensional, arbitrarily shaped, acoustically rigid body subjected to a plane wave incidence. The mathematical formulation is based on the potential theory and, in this work, the numerical method is applicable to the so-called double layer formulation (DLF). The scattering body is approximated by planar triangular patches. For the MoM solution using triangular patch modeling, edge-based basis functions are utilized to approximate the source distribution efficiently. These basis functions along with an appropriate testing procedure generates a simple numerical algorithm that is versatile enough to be applicable to closed, as well as open bodies. Finally, the present solution method is validated with several representative examples. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1703536]

PACS numbers: 43.20.Fn [LLT] Pages: 1926–1933

I. INTRODUCTION

In the last three decades, the problem of developing efficient numerical techniques to obtain the acoustic scattered/radiated fields from complex, three-dimensional bodies received considerable attention. This is partly due to the developments in digital computer technology and also partly due to the ever-increasing demand in the defense/commercial sector to obtain more faster/accurate field predictions. As a result, several formulations/algorithms have been developed based on the T-matrix approach,[1–4] boundary integral equation method,[5–8] and the method of moments (MoM).[9–14]

Although all these formulations work well with varying degrees of complexity, all of them break down at certain frequencies of the incident field. These frequencies happen to coincide with the resonance frequency of the cavity formed by the body. Several remedies have been proposed to overcome this problem. Notable among them are (a) the Combined Helmholtz Integral Equation Formulation (CHIEF) procedure[15] and (b) the Burton and Miller (BM) procedure.[16] The CHIEF method, although very popular, is somewhat heuristic and prone to inaccuracies, especially at high frequencies. On the other hand, the BM solution is more elegant and mathematically guaranteed to be stable.

The BM procedure basically suggests developing and solving two separate formulations known as the single layer formulation (SLF) and the double layer formulation (DLF). Burton and Miller proved mathematically that by combining SLF and DLF with appropriate constants, the breakdown phenomenon can be completely eliminated.[16] Here we note that SLF is relatively simple to solve,[9] whereas it is much more difficult to generate a general purpose solution based on DLF because of the hypersingular nature of the kernel in the integral equation. Although a few authors have attempted to generate numerical procedures by solving DLF,[17–19] these are often difficult to implement in a general purpose code.

At this stage, it is important to note that the breakdown phenomenon, as presented so far, is applicable only for closed bodies. Although, in a practical sense, everybody is a closed body, it is much more efficient to treat several geometrical shapes as open bodies by ignoring the finite thickness. For example, thin plates, disks and aircraft wings may be treated as open bodies, to name a few, because the thickness of such bodies is very small compared to other geometrical dimensions. However, all the available formulations, as far as our knowledge is concerned, treat the scattering body as a closed body. One of our main objectives of this work is to develop a unified solution where the scattering body may be either open or closed.

In the present work, we present a simple numerical method to solve the DLF via a well-known MoM procedure.[20] We note that the MoM procedure allows greater flexibility while approximating the unknown source distribution. We also note that the present solution scheme neither regularizes the hypersingular kernel nor implements special/complicated integration schemes. The advantages of the present scheme may be enumerated as follows.

(1) The DLF is used as is without any approximations/modifications either to the unknown function or to the kernel function.
(2) The numerical procedure is simple and requires no complicated extraction of singularities.
(3) The numerical procedures to evaluate the integrals are straightforward and efficient. Standard Gaussian integration schemes suffice to get an accurate solution.
(4) The procedure is very efficient when thin bodies are treated as open bodies. For this case, the number of unknowns required is drastically reduced.

Last, we note that although the SLF is used extensively, we believe that the DLF is more general in the sense that this

[a)]Corresponding author: Sadasiva M. Rao, Supercomputer Education and Research Center, Indian Institute of Science, Bangalore—560 012, India. Electronic mail: smrao@serc.iisc.ernet.in; phone: 91-80-360 0086; fax: 91-80-360 2648.

formulation is applicable to both open and closed bodies, as demonstrated in the following sections.

This paper is organized as follows.

In the next section, we briefly describe the MoM in general terms for the sake of completeness. In Sec. III, we describe the mathematical derivation of DLF. Next, in Sec. IV we describe the numerical solution procedure using MoM. Several representative examples are presented to illustrate the efficacy of the present approach in Sec. V. Last, in Sec. VI, we present some important conclusions drawn from this work.

II. OUTLINE OF THE METHOD OF MOMENTS

Consider the deterministic equation,

$$\mathbf{Lf}=\mathbf{g}, \tag{1}$$

where **L** is a linear operator, **g** is a known function, and **f** is an unknown function to be determined. Let **f** be represented by a set of known functions $\mathbf{f}_j$, $j=1,2,...,N$, termed as basis functions in the domain of **L** as a linear combination, given by

$$\mathbf{f}=\sum_{j=1}^{N} \alpha_j \mathbf{f}_j, \tag{2}$$

where α_j are scalar coefficients to be determined. Substituting Eq. (2) into Eq. (1), and using the linearity of **L**, we have

$$\sum_{j=1}^{N} \alpha_j \mathbf{L}\mathbf{f}_j=\mathbf{g}, \tag{3}$$

where the equality is usually approximate. Let $(\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3,\ldots)$ define a set of testing functions in the range of **L**. Now, taking the inner product of Eq. (3) with each $\mathbf{w}_i$ and using the linearity of inner product defined as $\langle \mathbf{f},\mathbf{g}\rangle = \int_s \mathbf{f}\cdot\mathbf{g}\, ds$, we obtain a set of linear equations, given by

$$\sum_{j=1}^{N} \alpha_j \langle \mathbf{w}_i, \mathbf{L}\mathbf{f}_j\rangle = \langle \mathbf{w}_i, \mathbf{g}\rangle, \quad i=1,2,3...,N. \tag{4}$$

The set of equations in Eq. (4) may be written in the matrix form as

$$\mathbf{ZX}=\mathbf{Y}, \tag{5}$$

which can be solved for **X** using any standard linear equation solution methodologies. The simplicity, accuracy, and efficiency of the method of moments lies in choosing the proper set of basis/testing functions and applying it to the problem at hand. In this work, we propose a special set of basis functions and a novel testing scheme to obtain accurate results using DLF.

III. MATHEMATICAL FORMULATION OF DLF

Let us consider an arbitrarily shaped three-dimensional acoustically rigid body surrounded by an infinite, homogeneous, nonviscous medium, such as air, as shown schematically in Fig. 1. Let S represent the surface of the body, and ρ and c be the density and speed of sound in the surrounding medium, respectively. Let $(p^i,\mathbf{u}^i)$ and $(p^s,\mathbf{u}^s)$ represent the incident and scattered pressure and velocity fields, respectively. It is important to note that the incident fields are defined in the absence of the scatterer. It is customary to introduce a velocity potential Φ such that $\mathbf{u}=\nabla\Phi$ and $p=-j\omega\rho\Phi$, assuming harmonic time variation.

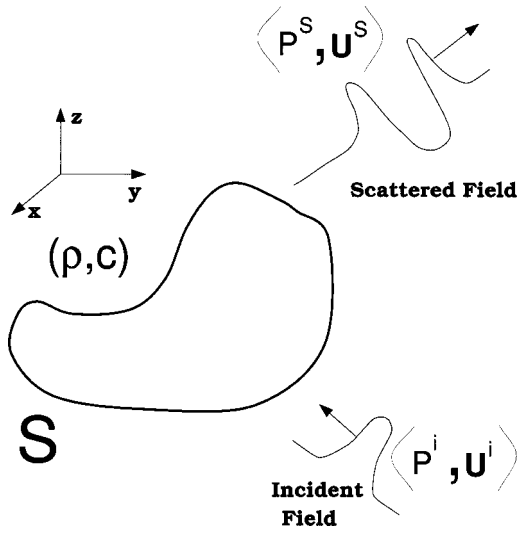


FIG. 1. Arbitrary body excited by an acoustic plane wave.

Let σ represent the simple source distribution on the scatterer. Using the potential theory and the free space Green's function, the scattered velocity potential may be defined as

$$\Phi^s=\int_S \sigma(\mathbf{r}')G(\mathbf{r},\mathbf{r}')ds' \tag{6}$$

and

$$\Phi^s=\int_S \sigma(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}ds', \tag{7}$$

for the SLF and DLF, respectively. In Eqs. (6) and (7),

$$G(\mathbf{r},\mathbf{r}')=\frac{e^{-jkR}}{4\pi R}, \tag{8}$$

and

$$R=|\mathbf{r}-\mathbf{r}'|; \tag{9}$$

$\mathbf{r}'$, $\mathbf{r}$, and k represent the locations of the source point, location of the observation point, and the wave number, respectively. Both $\mathbf{r}$ and $\mathbf{r}'$ are defined with respect to a global coordinate origin O. Also, note that in Eq. (7), $\partial/\partial n'$ represents the normal derivative with respect to the source point $\mathbf{r}'$. Noting that, at the surface of the hard scatterer, the normal derivative of the total velocity potential, which is the sum of the incident and scattered potentials, must vanish, the SLF and DLF integral equations may be derived, given by

$$\int_S \sigma(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n}ds' = -\frac{\partial \Phi^i(\mathbf{r})}{\partial n}, \tag{10}$$

and

$$\frac{\partial}{\partial n}\int_S \sigma(\mathbf{r}')\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}ds' = -\frac{\partial \Phi^i(\mathbf{r})}{\partial n}, \tag{11}$$

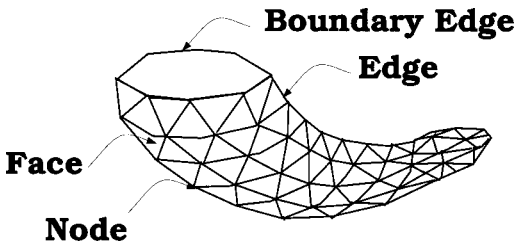


FIG. 2. Arbitrary body modeled by triangular patches.

respectively. In Eqs. (10) and (11), $\partial/\partial n$ represents the normal derivative with respect to the observation point $\mathbf{r}$. Also in Eqs. (10) and (11), Φ^i represents the incident velocity potential that is related to the incident pressure field p^i by the relation $p^i = -j\omega\rho\Phi^i$.

Next, we note that the SLF, given in Eq. (10), is relatively easy to solve,[9] whereas the DLF presented in Eq. (11) is much more complicated because of the double derivatives present in the integral equation. Before we attempt the numerical solution of DLF, we note that, after following the procedures developed in Refs. 21, 22, Eq. (11) may be rewritten as

$$\int_s \mathbf{a}_n \cdot \mathbf{a}_n' k^2 \sigma G \, ds' + \int_S (\mathbf{a}_n' \times \nabla' \sigma) \cdot (\mathbf{a}_n \times \nabla G) ds' = \mathbf{a}_n \cdot \nabla \Phi^i, \tag{12}$$

where $\mathbf{a}_n$ and $\mathbf{a}_n'$ represent the unit normal vectors at $\mathbf{r}$ and $\mathbf{r}'$, respectively. Next, for the sake of clarity, let us define

$$\mathbf{J} = \mathbf{a}_n \times \nabla \sigma. \tag{13}$$

Then, using Eq. (13), we can write Eq. (12) as

$$\int_S \mathbf{a}_n \cdot \mathbf{a}_n' k^2 \sigma G \, ds' + \int_S \mathbf{J} \cdot (\mathbf{a}_n \times \nabla G) ds' = \mathbf{a}_n \cdot \nabla \Phi^i. \tag{14}$$

In the next section, we develop the numerical solution to solve Eq. (14) and note that the unknown quantities to be evaluated are σ and $\mathbf{J}$. Further, we also note that σ and $\mathbf{J}$ are not independent quantities but related to each other by Eq. (13).

IV. NUMERICAL SOLUTION PROCEDURE

In this section, we present, in detail, the numerical solution of Eq. (14) for arbitrarily shaped scatterers. First of all, we approximate the arbitrary body using planar triangular patch modeling. The triangular patch modeling is the most efficient way of describing an arbitrary body to the digital computer. Further, there exists a variety of mesh generation codes that can be effectively used for the present numerical scheme.

Let the given arbitrary body be represented by triangular patches, as shown in Fig. 2. Further, let the triangular patch model consists of N_n nodes, N_e edges, and N_f faces. Also, let N_b represents the number of boundary edges if the body is open. We note that in the following solution scheme the number of unknowns N, i.e., the dimension of the moment matrix, is equal to $N_e - N_b$. Next, we define the basis functions to represent the unknown source distribution as required by the MoM procedure.

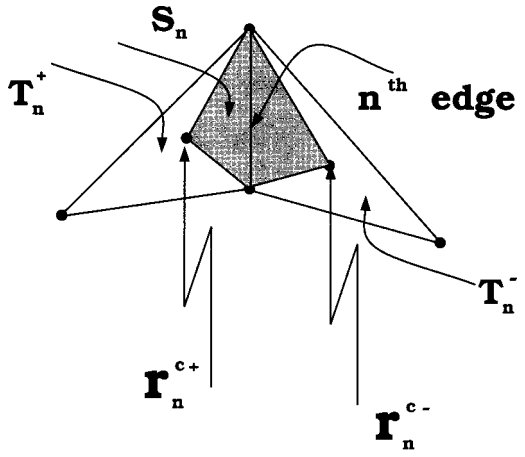


FIG. 3. Geometrical parameters associated with an interior edge.

A. Description of basis functions

It is convenient to start our development of basis functions to represent the unknown quantity σ by noting that each basis function is to be associated with an interior edge [i.e., nonboundary edge (cf. Fig. 2)] of the patch model. Let T_n^+ and T_n^- represent two triangles connected to edge n of the triangulated surface model, as shown in Fig. 3. The plus or minus designation of the triangles is arbitrary and of no consequence in the numerical solution scheme. We define the basis function associated with the nth edge as

$$f_n(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \in S_n; \\ 0, & \text{otherwise}; \end{cases} \tag{15}$$

where S_n represents the region obtained by connecting the centroids of triangles $T_n^{\pm}$, denoted by position vectors $\mathbf{r}_n^{c\pm}$, to the nodes of edge n. Note that this area is shown shaded in the Fig. 3. Using these basis functions, the unknown source distribution σ may be approximated as

$$\sigma(\mathbf{r}) = \sum_{n=1}^{N} \alpha_n f_n(\mathbf{r}), \tag{16}$$

where α_n represents the unknown coefficient to be determined.

Next, we define the basis functions to represent $\mathbf{J}$ as

$$\mathbf{J}(\mathbf{r}) = \sum_{n=1}^{N_f} \mathbf{g}_n(\mathbf{r}), \tag{17}$$

where $\mathbf{g}_n(\mathbf{r})$ is defined over each triangular patch and assumed to be constant over the patch. Since $\mathbf{g}_n$ is not an independent quantity, a relationship between α_n, f_n, and $\mathbf{g}_n$ may be derived as follows:

Consider a triangular patch T_p with associated nonboundary edges, p_1, p_2, and p_3, as shown in the Fig. 4. Then, using well-known Stoke's theorem and simple vector calculus, Eq. (13) may be rewritten as

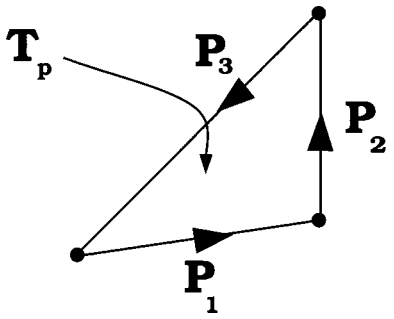


FIG. 4. Triangle T_p and associated edges.

$$\int_{T_p} \mathbf{J}\, ds = \int_{T_p} \mathbf{a}_n \times \nabla \sigma\, ds = \oint_{C_p} \sigma\, \mathbf{dl} = \alpha_{p_1} f_{p_1} \mathbf{l}_{p_1} + \alpha_{p_2} f_{p_2} \mathbf{l}_{p_2} + \alpha_{p_3} f_{p_3} \mathbf{l}_{p_3}, \tag{18}$$

where C_p is the contour bounded by the triangle T_p, and $\mathbf{l}_{p_i}$, $i=1, 2, 3$ represents the edge vectors, as shown in Fig. 4. We also note that

$$\int_{T_p} \mathbf{J}\, ds = \int_{T_p} \mathbf{g}_n\, ds = \mathbf{g}_n A_p, \tag{19}$$

where A_p represents the area of the triangle T_p. Last, using Eqs. (18) and (19), we have

$$\mathbf{g}_n = \frac{\alpha_{p_1} f_{p_1} \mathbf{l}_{p_1} + \alpha_{p_2} f_{p_2} \mathbf{l}_{p_2} + \alpha_{p_3} f_{p_3} \mathbf{l}_{p_3}}{A_p}. \tag{20}$$

The basis functions f_n, defined in Eq. (15), are uniquely suitable to solve the DLF. In the following, we note some advantages of the basis functions f_n.

(1) The areas S_n, $n=1,2,...,N$ represent nonoverlapping regions. Hence, it is easy to prove that the functions f_n are linearly independent.
(2) There is no unknown associated with the boundary edge. Thus, the source function is assumed to be zero on the boundary.
(3) The constant nature of the basis function helps in approximating σ to a high degree of accuracy. Although, it is generally felt that a basis with a linear or quadratic variation provide better approximation, it may not be true for an arbitrary distribution. What is true is if the nature of the distribution is known *a priori*, then one can design the basis functions that incorporate this variation and obtain an efficient solution. In the absence of such information, constant basis functions reasonably accurate solutions with less mathematical complexity. This point is well supported with several numerical examples in this work.
(4) The basis function can be easily extended to more complex geometries involving intersecting surfaces. Note that an intersecting surface, when modeled with triangular patches, may have more than two triangles connected to an edge. However, we are not considering the intersecting surfaces in this work since we are dealing with simple bodies only.
(5) Last, the functions f_n are helpful in dealing with the derivative operations in Eq. (14), as shown in the following sections.

The next step in the MoM is to select a testing procedure. For the sake of simplicity, we use the expansion functions defined in Eq. (15) also as testing functions. We note here that such a procedure is commonly known as the Galerkin procedure and results in a symmetric moment matrix provided no further approximations are made.[12] However, in the present procedure, for the sake of efficient evaluation of the matrix elements, we approximate the integrals involved in the testing procedure by a simple one-point integration rule, which, incidentally, destroys the symmetric nature of the matrix, and hence the procedure is not strictly the Galerkin procedure. Although not included here, one can perform higher-order integration and obtain a symmetric matrix. Obviously, when the system matrix is very large, symmetry property can be exploited to reduce the storage requirements.

B. Testing procedure

The testing procedure begins with defining the symmetric product as

$$\langle f_1, f_2 \rangle = \int_S f_1 f_2\, ds, \tag{21}$$

where f_1 and f_2 are two scalar functions defined over a surface S. Thus, the testing equation may be written as

$$\left\langle f_m, \int_S \mathbf{a}_n \cdot \mathbf{a}_n' k^2 \sigma G\, ds' \right\rangle + \left\langle f_m, \int_S \mathbf{J} \cdot (\mathbf{a}_n \times \nabla G) ds' \right\rangle = \langle f_m, \mathbf{a}_n \cdot \nabla \Phi^i \rangle. \tag{22}$$

We note that

$$\left\langle f_m, \int_S \mathbf{a}_n \cdot \mathbf{a}_n' k^2 \sigma G\, ds' \right\rangle = \int_S f_m(\mathbf{r}) \int_S \mathbf{a}_n \cdot \mathbf{a}_n' k^2 \sigma(\mathbf{r}') G(\mathbf{r}, \mathbf{r}') ds'\, ds \approx \left(\frac{A_m^+}{3} \mathbf{a}_m^+ + \frac{A_m^-}{3} \mathbf{a}_m^- \right) \cdot \int_S \mathbf{a}_n' k^2 \sigma(\mathbf{r}') G(\mathbf{r}_m, \mathbf{r}') ds', \tag{23}$$

where $A_m^{\pm}$ and $\mathbf{a}_m^{\pm}$ represent the areas and unit normal vectors of the triangles connected to the mth edge, respectively, and $\mathbf{r}_m$ represents the position vector to the midpoint of the mth edge. We also note that the double surface integration in Eq. (23) is converted to single surface integral by approximating the integrand at the center of the mth edge and multiplying by the area of the subdomain patch. This approximation is justified because of the following reasons.

(i) The subdomains are sufficiently small, which is a necessary requirement to obtain an accurate solution using the MoM.

(ii) The quantity being approximated undergoes an integration operation that makes it a reasonably smooth varying function over the region of integration.

Next, using the same logic and assuming the incident field to be a slowly varying function, we approximate the right-hand side of Eq. (22) as

$$\langle f_m, \mathbf{a}_n \cdot \nabla \Phi^i \rangle = \int_S f_m(\mathbf{r}) \mathbf{a}_n \cdot \nabla \Phi^i(\mathbf{r}) ds \approx \left(\frac{A_m^+}{3} \mathbf{a}_m^+ + \frac{A_m^-}{3} \mathbf{a}_m^- \right) \cdot \nabla \Phi^i(\mathbf{r}_m). \quad (24)$$

Now, we consider the evaluation of the $\langle f_m, \int_S \mathbf{J} \cdot (\mathbf{a}_n \times \nabla G) ds' \rangle$ term in Eq. (22).

For the sake of clarity, let us define

$$\mathbf{A} = \int_S \mathbf{J} G \, ds'. \quad (25)$$

Then, we have

$$\mathbf{J} \cdot \mathbf{a}_n \times \nabla G = \mathbf{a}_n \cdot \nabla G \times \mathbf{J} = \mathbf{a}_n \cdot [\nabla \times (G\mathbf{J}) - G \nabla \times \mathbf{J}]. \quad (26)$$

Here, we note that the second term on the right-hand side of Eq. (26) vanishes since the curl operation is on the unprimed variables only. Thus, we have

$$\int_S \mathbf{J} \cdot \mathbf{a}_n \times \nabla G \, ds' = \mathbf{a}_n \cdot \nabla \times \int_S \mathbf{J} G \, ds' = \mathbf{a}_n \cdot \nabla \times \mathbf{A}. \quad (27)$$

Using Eqs. (25)–(27), we can write

$$\left\langle f_m, \int_S \mathbf{J} \cdot (\mathbf{a}_n \times \nabla G) ds' \right\rangle = \int_S \mathbf{a}_n \cdot \nabla \times \mathbf{A} \, ds = \oint_{C_m} \mathbf{A} \cdot \mathbf{dl} \approx \mathbf{l}_m \cdot [\mathbf{A}(\mathbf{r}_m^{c-}) - \mathbf{A}(\mathbf{r}_m^{c+})], \quad (28)$$

where C_m and $\mathbf{l}_m$ are the contour of the basis function and the vector along the length associated with the mth edge, respectively.

Finally, using Eqs. (23), (24), and (28), the testing equation may be written as

$$\left(\frac{A_m^+}{3} \mathbf{a}_m^+ + \frac{A_m^-}{3} \mathbf{a}_m^- \right) \cdot \int_S \mathbf{a}_n' k^2 \sigma(\mathbf{r}') G(\mathbf{r}_m, \mathbf{r}') ds' + \mathbf{l}_m \cdot [\mathbf{A}(\mathbf{r}_m^{c-}) - \mathbf{A}(\mathbf{r}_m^{c+})] = \left(\frac{A_m^+}{3} \mathbf{a}_m^+ + \frac{A_m^-}{3} \mathbf{a}_m^- \right) \cdot \nabla \Phi^i(\mathbf{r}_m), \quad (29)$$

for $m = 1,2,...,N$.

C. Evaluation of matrix elements

Substituting the source expansion, Eqs. (16) and (17), into Eq. (29) yields an $N \times N$ system of linear equations that may be written in matrix form as

$$\mathbf{ZX} = \mathbf{Y}, \quad (30)$$

where $\mathbf{Z} = [Z_{mn}]$ is an $N \times N$ matrix and $\mathbf{X} = [\alpha_n]$ and $\mathbf{Y} = [Y_m]$ are column vectors of length N. Elements of $\mathbf{Z}$ and $\mathbf{Y}$ are given by

$$Z_{mn} = k^2 \left(\frac{A_m^+}{3} \mathbf{a}_m^+ + \frac{A_m^-}{3} \mathbf{a}_m^- \right) \cdot \mathbf{f}_{mn} + \mathbf{l}_m \cdot [\mathbf{A}_{mn}^{c-} - \mathbf{A}_{mn}^{c+}], \quad (31)$$

$$Y_m = \left(\frac{A_m^+}{3} \mathbf{a}_m^+ + \frac{A_m^-}{3} \mathbf{a}_m^- \right) \cdot \nabla \Phi^i(\mathbf{r}_m), \quad (32)$$

where

$$\mathbf{f}_{mn} = \mathbf{a}_n^+ \int_{S_n^+} \frac{e^{-jkR_m}}{4\pi R_m} ds' + \mathbf{a}_n^- \int_{S_n^-} \frac{e^{-jkR_m}}{4\pi R_m} ds', \quad (33)$$

$$\mathbf{A}_{mn}^{c\pm} = \mathbf{l}_n \left[\frac{1}{A_n^+} \int_{T_n^+} \frac{e^{-jkR_m^{c\pm}}}{4\pi R_m^{c\pm}} ds' + \frac{1}{A_n^-} \int_{T_n^-} \frac{e^{-jkR_m^{c\pm}}}{4\pi R_m^{c\pm}} ds' \right], \quad (34)$$

$$R_m = |\mathbf{r}_m - \mathbf{r}'|, \quad (35)$$

$$R_m^{c\pm} = |\mathbf{r}_m^{c\pm} - \mathbf{r}'|. \quad (36)$$

Integrals, appearing in Eqs. (33) and (34), are straightforward integrals over a triangular region. However, it is cautioned that the integrals have singular kernels and, for an accurate solution, may be evaluated using the methods described in Refs. 23, 24.

For the plane wave incidence, we set

$$\Phi^i = e^{jk\mathbf{k} \cdot \mathbf{r}}, \quad (37)$$

where the propagation vector $\mathbf{k}$ is

$$\mathbf{k} = \sin\theta_0 \cos\phi_0 \mathbf{a}_x + \sin\theta_0 \sin\phi_0 \mathbf{a}_y + \cos\theta_0 \mathbf{a}_z, \quad (38)$$

and (θ_0, ϕ_0) defines the angle of arrival of the plane wave in terms of the usual spherical coordinate convention.

Once the elements of the moment matrix $\mathbf{Z}$ and the forcing vector $\mathbf{Y}$ are determined, one may solve the resulting system of linear equations, Eq. (30), for the unknown column vector $\mathbf{X}$.

V. NUMERICAL RESULTS

In this section, numerical results are presented for selected scatterers under plane wave excitation. The geometries considered are a sphere, a cube, a finite cylinder, a cone, a square plate, and a circular disk. For all cases, the body is placed at the center of the coordinate system and the plane wave is traveling along the Z axis with wave number $k = 1$. It may be noted that although several different examples have been tested, we are presenting only some representative cases for validation. Further, no convergence studies have been carried out to ascertain the optimum number of patches required to obtain a certain percentage of accuracy. In all cases the scattering cross section S is given by

$$S = 4\pi r^2 \left| \frac{\Phi^s}{\Phi^i} \right|^2. \quad (39)$$

For comparison, we also present the results obtained using the SLF as described in Ref. 9.

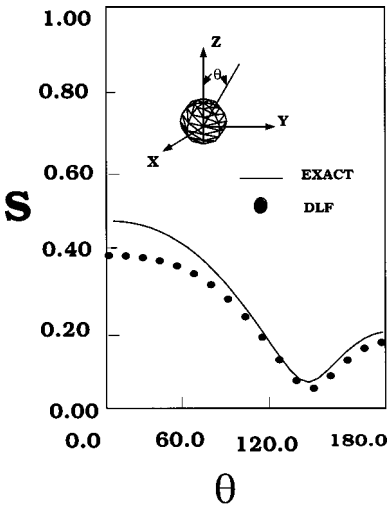


FIG. 5. Scattering cross section versus polar angle for an acoustically hard sphere, radius (a)=1 m, $ka=1.0$, subjected to an axially incident plane wave.

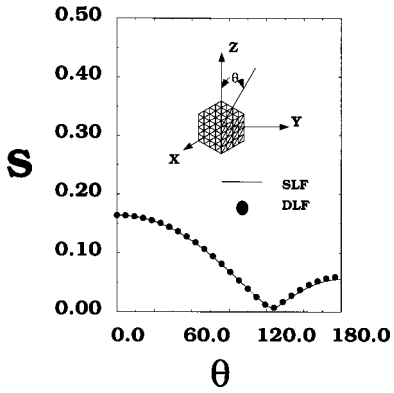


FIG. 7. Scattering cross section versus polar angle for an acoustically hard cube, length=1 m, subjected to an axially incident plane wave.

As a first example, consider a sphere of radius 1 m. The sphere is an important geometry because of its doubly curved surface, which is not amenable to rectangular patches. Further, the triangular patch modeling of a spherical geometry involves nonuniform triangles that test the numerical procedure to a full extent. For numerical purposes, the sphere is modeled by 112 triangular patches and 168 edges. The modeling is done by first dividing the θ and ϕ directions into eight equal segments each. This results in triangular patches at the sphere caps and quadrilateral patches in between. By joining the diagonal of the quadrilateral patch, the triangular discretization may be obtained. Figure 5 shows the scattering cross section S as a function of the polar angle θ for the hard sphere. It is evident from the figure that the method presented in this paper compares reasonably well with the exact solution.[25] The small discrepancy between the two solutions can be attributed to an insufficient number of unknowns and the fact that the triangulated sphere has a lower surface area than the actual sphere.

In order to demonstrate the convergence of the numerical method, we present the scattering cross section S as a function of the polar angle θ for the hard sphere ($ka=2.0$) in Fig. 6. Here, we compare the numerical solution obtained using $N=168$, 462, and 918, with the exact solution. It is clear from the figure that the solutions for $N=462$ and 918 compare very well with the exact solution, even though an expanded scale is adopted for plotting the figure. Also, there is little difference between the $N=432$ and 928 solutions, indicating the convergence of the solution.

Next, we consider the case of a cube with side length $l=1.0$ m. The case of a cube presents a challenging task of handling sharp edges and corners. To obtain a triangular patch model, each side of the cube is divided into 6 equal segments resulting in 216 square patches on the cube. By joining the diagonals, we get 432 triangular patches and 648 edges. Figure 7 shows the scattering cross section S as a function of θ. It is evident from the figure that the constant basis functions defined over an edge for DLF compare very well with that of SLF. Also, note that even though the basis functions are defined on sharp edges and corners, the numerical results are fairly accurate. Finally, we note that the node-based conventional boundary integral methods[5–8] need complex calculations to obtain accurate results for this geometry.

As a third example, we consider the case of a finite cylinder of height 2.0 and 1.0 m radius. The triangular patch modeling of the cylinder is obtained by dividing the length and circumference into 8 and 10 uniform segments, respectively, resulting in 80 rectangular patches. By joining the diagonals, we obtain 160 triangular patches. The cylinder is closed on both ends by circular disks that are modeled by

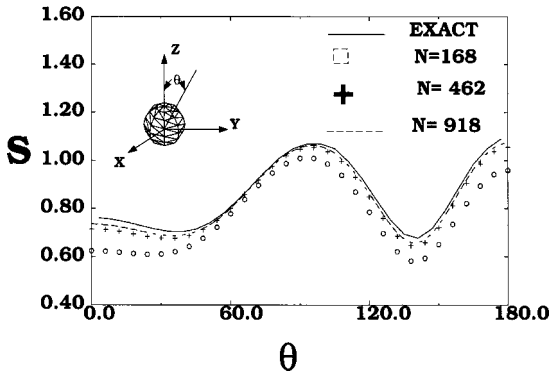


FIG. 6. Scattering cross section versus polar angle for an acoustically hard sphere, radius (a)=1 m, $ka=2.0$, subjected to an axially incident plane wave.

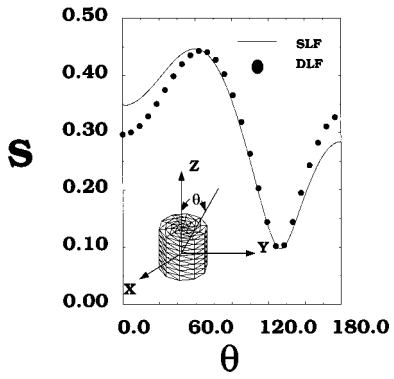


FIG. 8. Scattering cross section versus polar angle for an acoustically hard cylinder, radius=1 m and height=2 m, subjected to an axially incident plane wave.

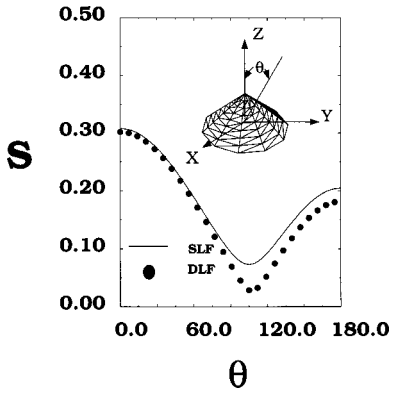


FIG. 9. Scattering cross section versus polar angle for an acoustically hard cone, radius=1 m and height=1 m, subjected to an axially incident plane wave.

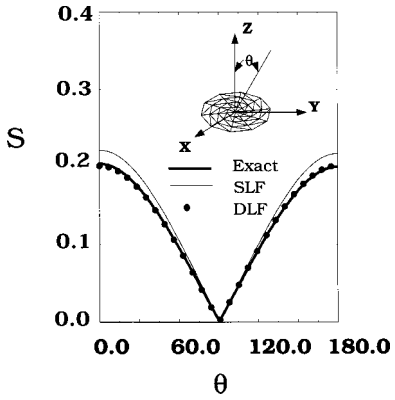


FIG. 11. Scattering cross section versus polar angle for a rigid circular plate, radius=1 m, subjected to an axially incident plane wave.

and additional 70 triangular patches each. Thus, in total we have 300 patches and 450 edges for this geometry. Figure 8 shows the scattering cross section S as a function of θ. For this case also, we note a good comparison for SLF and DLF results.

Next, we consider a cone of 1 m height and 1 m radius. The cone has a sharp corner at the tip that is difficult to handle using node-based BIE methods. Using a similar discretization scheme as in the case of a finite cylinder, the cone is divided into 220 patches resulting in 330 edges. Figure 9 shows the scattering cross section S as a function of θ, which compares very well with the SLF result.

Now we consider two more scatterers, viz., a square plate and a circular disk, whose thickness is very small compared to the other dimensions. Here we note that, for such type of problems, the conventional SLF requires careful modeling to obtain accurate results. Since the thickness is small, one needs to use small triangles over the entire body consisting of two identical surfaces and sidewalls in order to have a more or less uniform triangularization. This requirement results in a very large number of triangles for the solution scheme, which is computationally very expensive. However, these type of problems can be handled very easily using the present formulation by approximating the scattering body as an open structure with zero thickness. Since we are modeling only one surface, the number of unknowns for DLF is very low, as is evident from the following examples.

Consider a square plate of length 1.0 m and thickness 0.01 m, placed in the XY plane, with the center of the plate coinciding with the coordinate origin, as shown in the inset of Fig. 10 and excited by an acoustic plane wave traveling along the Z axis. Using the DLF method and assuming the scatterer as an open structure, the square plate is modeled by 32 triangular patches with 40 nonboundary edges. Thus, for this case, the dimension of the moment matrix is 40. The solution obtained by the DLF scheme is compared with the SLF method and presented in Fig. 10. For the SLF method, we have modeled the plate by dividing each side and the thickness into 20 divisions and 1 division, respectively, resulting in 1760 triangular patches. We note that the comparison between the two methods is quite acceptable. We also note that, using an even finer mesh scheme for SLF, one can obtain a better comparison. It is obvious that the DLF method is much more efficient for this case and converges to a stable result with very few unknowns.

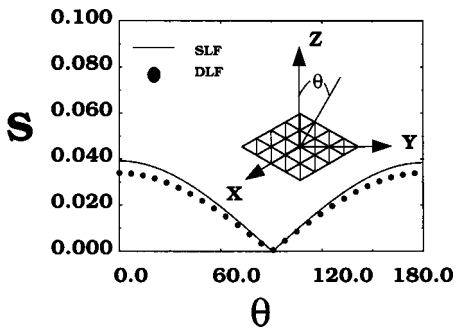


FIG. 10. Scattering cross section versus polar angle for a rigid square plate, side length=1 m, subjected to an axially incident plane wave.

As a last example, we consider a circular disk of radius 1 m and thickness 0.05 m with the center coinciding with the origin of the coordinate system and lying in the XY plane, as shown in the inset of Fig. 11. The disk is excited by an acoustic plane wave traveling along the Z axis. The circular disk is modeled by 70 triangular patches and 100 nonboundary edges for the DLF method. For the SLF scheme, we have used 2400 triangles. The scattering cross section is presented for both cases in Fig. 11 and, for comparison, we also present the exact solution.[25] We note a very good comparison between the exact and DLF results. We observe that the discrepancy in the SLF results is due to the disk being not very thin. Last, we believe that the SLF procedure becomes prohibitively expensive, in terms of computational resources, for even thinner disks than considered here and the DLF generates stable and accurate results for such situations.

VI. CONCLUSIONS

In this work, we have presented a numerical solution to the acoustic scattering problem by arbitrarily shaped, three-dimensional rigid bodies using the double layer formulation (DLF). The numerical solution is obtained via a method of moments solution procedure. In order to develop a simple solution, we have used planar triangular patch modeling to describe the arbitrary body and employed a special set of edge-based basis functions to approximate the unknown source distribution. One of the important conclusions that

can be inferred from this work is that the edge-based solution is less complicated than the node-based solution because of no complex singularity extraction in the integrals involved. Further, the DLF, as developed in this work, is more general than the so-called single layer formulation (SLF), in the sense that the present procedure is applicable to closed as well as open bodies. The open body solution results in a very efficient solution since the number of unknowns are markedly few compared to the closed body solution. We have demonstrated this fact using square plate and circular disk as examples. It is obvious that a complex structure, such as an aircraft, can be envisaged as a combination of both open and closed bodies and we believe that for such cases the present formulation generates the most efficient solution. Presently, work is in progress to apply the present technique to complex bodies.

Next, the work is also underway to eliminate problem on internal resonance by combining both SLF and DLF solutions.[16] Although the SLF solution is reported earlier,[9] it is actually a face-based solution that implies that the basis functions have been defined on the triangular face. However, to combine the present DLF solution with the SLF solution, we need an edge-based SLF solution that will be reported in due course.

Finally, we believe that the work presented in this paper is applicable to other acoustic field problems such as fluid bodies, rigid bodies coated with fluid materials, and elastic bodies.

[1] *Acoustic, Electromagnetic and Elastic Wave Scattering—Focus on the T-Matrix Approach*, edited by V. K. Varadan and V. V. Varadan (Pergamon, New York, 1980).

[2] V. V. Varadan, A. Lakathia, and V. K. Varadan, ''Comments on recent criticism of the T-Matrix method,'' J. Acoust. Soc. Am. **84**, 2280–2284 (1988).

[3] W. Tobocman, ''Calculation of acoustic wave scattering by means of the Helmholtz integral equation I,'' J. Acoust. Soc. Am. **76**, 595–607 (1984).

[4] W. Tobocman, ''Calculation of acoustic wave scattering by means of the Helmholtz integral equation II,'' J. Acoust. Soc. Am. **76**, 1549–1554 (1984).

[5] G. T. Schuster and L. C. Smith, ''A comparison and four direct boundary integral methods,'' J. Acoust. Soc. Am. **77**, 850–864 (1985).

[6] G. T. Schuster, ''A hybrid BIE+Born series modeling scheme: Generalized Born series,'' J. Acoust. Soc. Am. **77**, 865–879 (1985).

[7] A. F. Seybert, B. Soenarko, F. J. Rizzo, and D. J. Shippy, ''An advanced computation method for radiation and scattering acoustic waves in three dimensions,'' J. Acoust. Soc. Am. **77**, 362–368 (1985).

[8] P. Malbequi, S. M. Candel, and E. Rignot, ''Boundary integral calculations of scattered fields: Application of spacecraft launcher,'' J. Acoust. Soc. Am. **82**, 1771–1781 (1987).

[9] S. M. Rao and P. K. Raju, ''Application of Method of moments to acoustic scattering from multiple bodies of arbitrary shape,'' J. Acoust. Soc. Am. **86**, 1143–1148 (1989).

[10] P. K. Raju, S. M. Rao, and S. P. Sun, ''Application of the method of moments to acoustic scattering from multiple infinitely long fluid filled cylinders,'' Comput. Struct. **39**, 129–134 (1991).

[11] S. M. Rao and B. S. Sridhara, ''Application of the method of moments to acoustic scattering from arbitrary shaped rigid bodies coated with lossless, shearless materials of arbitrary thickness,'' J. Acoust. Soc. Am. **90**, 1601–1607 (1991).

[12] S. M. Rao, P. K. Raju, and S. P. Sun, ''Application of the method of moments to acoustic scattering from fluid-filled bodies of arbitrary shape,'' Commun. Appl. Numer. Methods **8**, 117–128 (1992).

[13] S. M. Rao and B. S. Sridhara, ''Acoustic scattering from arbitrarily shaped multiple bodies in half-space: Method of moments solution,'' J. Acoust. Soc. Am. **91**, 652–657 (1992).

[14] S. P. Sun and S. M. Rao, ''Application of the method of moments to acoustic scattering from multiple infinitely long fluid-filled cylinders using three different formulation,'' Comput. Struct. **43**, 1147–1153 (1992).

[15] H. A. Schenck, ''Improved integral formulation for acoustic radiation problems,'' J. Acoust. Soc. Am. **44**, 41–58 (1968).

[16] A. J. Burton and G. F. Miller, ''The application of integral equation methods to the numerical solution of some exterior boundary value problems,'' Proc. R. Soc. London, Ser. A **323**, 601–618 (1971).

[17] S. Amini and D. T. Wilton, ''An investigation of boundary element methods for the exterior acoustic problem,'' Comput. Methods Appl. Mech. Eng. **54**, 49–65 (1986).

[18] W. L. Meyer, W. A. Bell, B. T. Zinn, and M. P. Stallybras, ''Boundary integral solutions of three dimensional acoustic radiation problems,'' J. Sound Vib. **59**, 245–262 (1978).

[19] C. C. Chien, H. Raliyah, and S. N. Alturi, ''An effective method for solving the hyper singular integral equations in 3-D acoustics,'' J. Acoust. Soc. Am. **88**, 918–937 (1990).

[20] R. F. Harrington, *Field Computation by Method of Moments* (MacMillan, New York, 1968).

[21] A. W. Maue, ''Zur formulierung eines allgemeinen beugungsproblems durch eine integralgleichung,'' J. Phys. (Moscow) **126**, 601–618 (1949).

[22] K. M. Mitzner, ''Acoustic scattering from an interface between media of greatly different density,'' J. Math. Phys. **7**, 2053–2060 (1966).

[23] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Bulter, ''Potential Integrals for uniform and linear source distributions on polygons and polyhedra domains,'' IEEE Trans. Antennas Propag. **AP-32**, 276–281 (1984).

[24] P. C. Hammer, O. P. Marlowe, and A. H. Stroud, ''Numerical integration over simplexes and cones,'' Math. Tables Aids Comput. **10**, 130–138 (1956).

[25] J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, *Electromagnetic and Acoustic Scattering by Simple Shapes* (North-Holland, Amsterdam, 1969).

Rapid calculations of time-harmonic nearfield pressures produced by rectangular pistons

Robert J. McGough[a)]
Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan 48864

(Received 22 December 2003; revised 8 February 2004; accepted 9 February 2004)

A rapid method for calculating the nearfield pressure distribution generated by a rectangular piston is derived for time-harmonic excitations. This rapid approach improves the numerical performance relative to the impulse response with an equivalent integral expression that removes the numerical singularities caused by inverse trigonometric functions. The resulting errors are demonstrated in pressure field calculations using the time-harmonic impulse response solution for a rectangular source 5 wavelengths wide by 7.5 wavelengths high. Simulations using this source geometry show that the rapid method eliminates the singularities introduced by the impulse response. The results of pressure field computations are then evaluated in terms of relative errors and computational speeds. The results show that, when the same number of Gauss abscissas are applied to both approaches for time-harmonic pressure field calculations, the rapid method is consistently faster than the impulse response, and the rapid method consistently produces smaller maximum errors than the impulse response. For specified maximum error values of 10% and 1%, the rapid method is 2.6 times faster than the impulse response for pressure field calculations performed on a 61 by 101 point grid. The rapid approach achieves even greater reductions in the computation time for smaller errors and larger grids. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1694991]

PACS numbers: 43.20.Rz, 43.20.El, 43.35.Bf [ADP] Pages: 1934–1941

I. INTRODUCTION

The impulse response approach popularized by Stepanishen[1] and derived by Lockwood and Willette for a rectangular source[2] provides a general method for calculating the nearfield of uniformly excited acoustic radiators with exact closed-form expressions. This approach defines the impulse response for each spatial coordinate as the response to an impulse velocity evaluated across the surface of a vibrating piston. The expression for the impulse response is convolved with the time derivative of the piston velocity, and the result describes the pressure output as a function of time. For pressure fields produced by time-harmonic excitations, the simulated field is directly proportional to the Fourier transform of the impulse response. Impulse response solutions are available for transducers with a wide variety of shapes, including standard circular,[1] rectangular,[2,3] and spherical shell[4] geometries. The impulse response is also applicable to simulations of transducers with nonuniform surface excitations.[5,6]

Unfortunately, acoustic field computations with the impulse response sometimes encounter numerical difficulties.[7] These numerical problems arise in response to rapid changes in the impulse response in regions above the edge of the piston. Numerical problems in these regions are typically addressed with high sampling rates. By increasing the number of samples and compensating for the increased frequency content of the impulse response, the computation time is therefore also increased.

With a new rapid integral formulation, the numerical problems with impulse response calculations are solved for uniformly excited rectangular pistons. In time-harmonic numerical calculations, the edge artifacts are eliminated when analytically equivalent integrals are derived and singularities are subtracted from the equivalent integrals. The equivalent integral improves the numerical performance both in the neighborhood of the edge and throughout the nearfield. The improved performance is demonstrated in numerical calculations which show that the integrals with the subtracted singularities converge much more quickly than integrals that evaluate the impulse response. Comparisons between these two methods show that the impulse response produces larger numerical errors and requires more computation time. Thus, the rapid formulation simultaneously reduces the computation time and decreases the numerical error relative to the impulse response.

II. THEORY

The impulse response formulation for a rectangular radiator is derived from the time-domain Green's function analysis presented in Lockwood and Willette,[2] which defines the steady-state acoustic field produced by a rectangular radiator for a time-harmonic excitation as

$$p(x,y,z;t) = -j\omega\rho v e^{j\omega t} H(x,y,z;k). \tag{1}$$

In Eq. (1), ω is the excitation frequency in radians per second, ρ is the density of the medium, v is a constant normal velocity evaluated at the surface of the rectangular radiator, k is the wavenumber, and $H(x,y,z;k)$ is the Fourier transform of the impulse response. The center of the rectangular radiator is the origin of the spatial coordinates (x,y,z), and the positive z direction is defined by the normal evaluated at the center of the rectangular source. When computed in terms of

a)Electronic mail: mcgough@egr.msu.edu

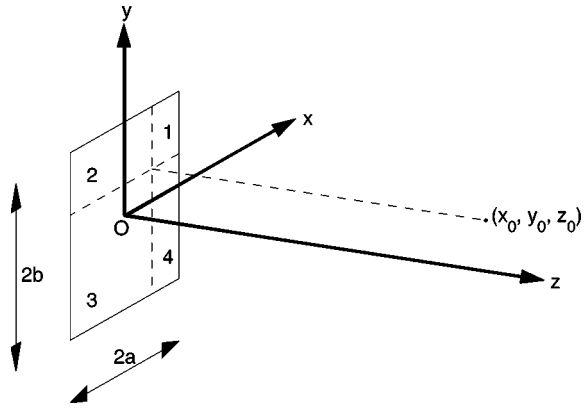


FIG. 1. Definition of coordinate axes and rectangles for superposition calculations. The center of the element defines the origin (O) of the coordinate system, and the vertices of the rectangular radiator are specified by the intersections of the lines $x=\pm a$ and $y=\pm b$. The z axis is coincident with the element normal. In this coordinate system, the borders of the rectangles defining the subelements required for superposition calculations are specified by the lines $x=x_0$, $y=y_0$, $x=\pm a$, and $y=\pm b$. All of these lines are located in the $z=0$ plane. The rectangles defined by these boundaries share a common corner at (x_0,y_0), and the Fourier transform of the impulse response at (x_0,y_0,z_0) is determined by superposing the results for rectangles that share the vertex (x_0,y_0).

the wavenumber k for a rectangular radiator with lateral dimensions $s\times l$, the Fourier transform of the impulse response evaluated directly above one corner is

$$H_{s,l}(z;k)=-\frac{1}{2\pi}\left(\frac{j\pi}{2k}[e^{-jk\sqrt{z^2+s^2+l^2}}-e^{-jkz}]\right.$$
$$-\int_{\sqrt{z^2+s^2}}^{\sqrt{z^2+s^2+l^2}}\cos^{-1}\left\{\frac{s}{\sqrt{\beta^2-z^2}}\right\}e^{-jk\beta}\,d\beta$$
$$\left.-\int_{\sqrt{z^2+l^2}}^{\sqrt{z^2+s^2+l^2}}\cos^{-1}\left\{\frac{l}{\sqrt{\beta^2-z^2}}\right\}e^{-jk\beta}\,d\beta\right). \tag{2}$$

After applying the change of variables $\tau=\beta/c$ and replacing the wavenumber k with ω/c, where c represents the speed of sound, this expression is equivalent to that presented in Lockwood and Willette.[2] In Eq. (2), the subscripts s and l contain the lengths of the lateral dimensions of the rectangular radiator measured in the x and y directions, respectively. The notation $H_{s,l}(z;k)$ emphasizes that the expression in Eq. (2) is only valid along the normal evaluated at the corner of each rectangular subelement. Therefore, this expression is only evaluated as a function of the z coordinate and the wavenumber k.

At all other field coordinates, the Fourier transform of the impulse response is determined by superposition according to the approach presented in Lockwood and Willette.[2] The superposition approach, which is adopted here, subdivides a rectangular radiator into four smaller rectangles whenever the (x,y) coordinates of the field point fall within the lateral (x,y) extent of a rectangular radiator that defines the z axis in the direction of the element normal. In Fig. 1, the source is divided into four smaller rectangles that share a corner at (x_0,y_0) for all points satisfying $|x|<a$ and $|y|<b$, where the element half-width is defined by a and the element half-height is defined by b. After the rectangular source is subdivided into subelements labeled 1, 2, 3, and 4, the sum of the individual contributions in front of the common corner point is then evaluated above the common corner of the four sources. Thus, the contributions of the four rectangular subelements are superposed and the Fourier transform of the total impulse response is evaluated as

$$H(x,y,z;k)=\sum_{i=1}^{4}\pm H_{s_i,l_i}(z;k). \tag{3}$$

In Eq. (3), the subscripts of s and l specify the subelement number as in Fig. 1. The sign of each contribution in Eq. (3) depends on the location of the field coordinate (x,y,z) relative to the outer boundary of the source. Four contributions are added together where $|x|<a$ and $|y|<b$, and, in general, when $|x|>a$ and/or $|y|>b$, two contributions are added and two more are subtracted. The exceptions to these rules occur wherever only one equality (either $|x|=a$ or $|y|=b$) is satisfied, in which case only two rectangular sources are superposed. Where $|x|=a$ and $|y|=b$, the field coordinate is located directly over a corner of the rectangular source so Eq. (2) is applied directly.

III. METHODS

The numerical performance of the integral presented in Eq. (2) is improved in three steps. The first improvement is obtained after equivalent integrals are derived and a singularity is subtracted from each integrand. The resulting equivalent analytical expression demonstrates superior numerical properties relative to the impulse response in terms of both computation time and numerical error. The second improvement further reduces the computation time by isolating repeated calculations that are unique to the new integral expression. Instead of evaluating the same expression repeatedly, values are computed once and then stored in memory for subsequent evaluations. The third step consolidates certain integrals that share the same integrand and a common upper or lower integration limit. After all of these steps are combined, the resulting integrals are evaluated with Gauss quadrature and normalized numerical errors are computed.

A. Equivalent integral formulation

When the integrals in Eq. (2) are evaluated numerically, the results are hampered by poor convergence characteristics. For these two integrals, the numerical convergence is limited by the slope of the integrand, which is dominated by the inverse cosine term in Eq. (2). The inverse cosine term produces slopes that are particularly large in all locations where $x\approx a$ and/or $y\approx b$. These large slopes cause considerable difficulty when standard quadrature techniques (trapezoidal rule, Simpson's rule, Newton–Cotes formulas, Romberg integration, Gauss quadrature, etc.) are employed. The numerical problems encountered by Eq. (2) motivate the derivation of an equivalent formulation that avoids excessive slopes within the integrand.

An equivalent integral is obtained from the first integral in Eq. (2) after $\cos^{-1}(s/\sqrt{\beta^2-z^2})$ is replaced with $\tan^{-1}(\sqrt{\beta^2-z^2-s^2}/s)$ and then the order of integration is exchanged as follows:

$$\int_{\sqrt{z^2+s^2}}^{\sqrt{z^2+s^2+l^2}} \cos^{-1}\left\{\frac{s}{\sqrt{\beta^2-z^2}}\right\} e^{-jk\beta}\,d\beta$$

$$= \int_{\sqrt{z^2+s^2}}^{\sqrt{z^2+s^2+l^2}} \left[\int_0^{\sqrt{\beta^2-z^2-s^2}} \frac{s}{\sigma^2+s^2}\,d\sigma\right] e^{-jk\beta}\,d\beta$$

$$= \int_0^{l} \left[\int_{\sqrt{\sigma^2+z^2+s^2}}^{\sqrt{l^2+z^2+s^2}} e^{-jk\beta}\,d\beta\right] \frac{s}{\sigma^2+s^2}\,d\sigma. \tag{4}$$

After the innermost integral is evaluated, the same procedure is repeated for the second integral in Eq. (2). The two resulting integral expressions are then inserted into Eq. (2), and the result is

$$H_{s,l}(z;k) = \frac{j}{2\pi k}\left(\frac{\pi}{2}e^{-jkz} - s\int_0^{l} \frac{e^{-jk\sqrt{z^2+\sigma^2+s^2}}}{\sigma^2+s^2}\,d\sigma - l\int_0^{s} \frac{e^{-jk\sqrt{z^2+\sigma^2+l^2}}}{\sigma^2+l^2}\,d\sigma\right). \tag{5}$$

Equation (5) is analytically equivalent to the integral expression in Eq. (2).

In Eq. (5), a singularity is encountered in each integrand whenever s or l approaches zero. The singularity arises if the corresponding integrand is evaluated at or near the lower limit. When the singularity is encountered in this location, the denominator becomes very small, and therefore the quotient becomes very large. This singularity is easily eliminated when a series expansion is evaluated for the numerator of each integrand.[8] The resulting term, e^{-jkz}, is then subtracted from and added to the numerator of each integrand. The term that is subtracted from each numerator is retained for numerical calculations, and the term that is added to each numerator is evaluated analytically. This yields

$$H_{s,l}(z;k) = -\frac{j}{2\pi k}\left(s\int_0^{l} \frac{e^{-jk\sqrt{z^2+\sigma^2+s^2}} - e^{-jkz}}{\sigma^2+s^2}\,d\sigma + l\int_0^{s} \frac{e^{-jk\sqrt{z^2+\sigma^2+l^2}} - e^{-jkz}}{\sigma^2+l^2}\,d\sigma\right). \tag{6}$$

Thus, an expression that is analytically equivalent to the impulse response is obtained, and by subtracting the singularity from the numerator of each integrand, the numerical conditioning of each integrand is further improved.

B. Shared integrands

In the region where $|x|>a$ and $|y|>b$, superposition dictates that contributions from two subelements are added and two subelements are subtracted according to Eq. (3). When these responses are superposed, integrals sharing a common integrand are combined to improve numerical performance. In this region, the pairs (s_1,l_1) and (s_2,l_2) are defined as the sides of the smallest and largest rectangles such that $s_1=|x|-a$, $l_1=|y|-b$, $s_2=|x|+a$, and $l_2=|y|+b$, where a is half of the element width and b is half of the element height defined in Fig. 1. The Fourier transform is then evaluated as $H(x,y,z;k)=H_{s_1,l_1}(z;k)+H_{s_2,l_2}(z;k)-H_{s_1,l_2}(z;k)-H_{s_2,l_1}(z;k)$. Ordinarily, this would require the evaluation of two integrals for each subelement, resulting in a total of eight integral evaluations; however, combining the limits for shared integrands reduces the number of integrals from eight to four. Two of the four integrals are evaluated from s_1 to s_2, and the two remaining integrals are evaluated from l_1 to l_2.

Combining integrals that share integrands reduces the total number of integrals evaluated, and this in turn decreases the computation time. The combined integrals also improve the accuracy of the computed acoustic field, since each integral is evaluated over a smaller range of values as specified by the limits. This approach is applicable whenever two terms share the same integrands, which occurs where $|x|>a$ and/or $|y|>b$. If only one of these two inequalities is satisfied, then two pairs of integrals are combined. The computation times and numerical errors are simultaneously reduced whenever either one or two pairs of integrals are combined.

C. Gauss quadrature

Each integral is evaluated numerically with Gauss quadrature.[8] The Gauss quadrature rule was obtained from the FORTRAN routine GRULE.[8] The Gauss rule computes the abscissas g_i and weights w_i of the n point Gauss–Legendre integration rule for the interval $[-1,1]$ in two iterations for double precision accuracy. The abscissas are then converted with a linear mapping function[9] that is defined for an arbitrary interval $[u,v]$. The mapping function is defined for the abscissas as

$$\sigma_i = \frac{v-u}{2}g_i + \frac{v+u}{2} \tag{7}$$

and the Gauss weights w_i are scaled by $(v-u)/2$. Gauss quadrature is applied in all of the simulation results presented here.

D. Repeated calculations

In Eq. (6), the limits of integration and the denominator of each integral are independent of the z coordinate. Furthermore, the same limits of integration are repeated in certain integrals, and certain expressions are repeatedly evaluated within the same integral. These features of Eq. (6) are exploited for numerical calculations with careful bookkeeping. If the field coordinate system is defined parallel to the planes where $x=0$ or $y=0$ (or both, as in the coordinate system of Fig. 1), the terms in the numerator and denominator containing σ^2+l^2 and/or σ^2+s^2 are computed once and then stored in memory for calculations using subsequent values of the z coordinate. This option is not available with the impulse response, which updates the limits of integration and recomputes the integrand with each new value of x, y, and z. Fur-

thermore, the abscissas appear in Eq. (6) only as square terms, and this is exploited by squaring the mapping function in Eq. (7), yielding

$$\sigma_i^2 = \frac{(v-u)^2}{4} g_i^2 + \frac{v^2-u^2}{2} g_i + \frac{(v+u)^2}{4}. \tag{8}$$

Thus, the squared Gauss abscissas g_i^2 are evaluated in advance, and the values for σ_i^2 are then determined from the linear combination of precomputed Gauss abscissas g_i and squared Gauss abscissas g_i^2. Additional simplifications are possible when the lower integration limit is zero, as in Eq. (6), allowing the reduction of Eq. (8) to $(v^2/4)(g_i+1)^2$, which is an expression that is further accelerated by precomputing $(g_i+1)^2$. Precomputed values of g_i^2 are available for both Eq. (2) and Eq. (6); however, in Eq. (2), the integrand contains both squared (β^2) and unsquared (β) abscissas, and the impulse response requires additional time to calculate both of these terms.

E. Error calculations

The spatial distribution of numerical errors in the computed acoustic field is obtained after the absolute value of the difference between two beam patterns is computed and the result is normalized. This error calculation, when presented in a mesh plot, highlights the regions where singularities are encountered. The error $\eta(n,z)$ describes the normalized difference between a complex pressure field, $P(x,z)$, and a reference beam pattern, $P_{\text{ref}}(x,z)$. The normalization factor is defined here as the maximum absolute value of the reference beam pattern $P_{\text{ref}}(x,z)$. This scalar normalization factor was selected to prevent division by zero and to avoid exaggerating the error values where the field amplitudes are relatively small. The spatially varying error $\eta(x,z)$ is thus defined as

$$\eta(x,z) = \frac{|P(x,z) - P_{\text{ref}}(x,z)|}{\max_{x,z}|P_{\text{ref}}(x,z)|}, \tag{9}$$

and the maximum error is then

$$\eta_{\max} = \frac{\max_{x,z}|P(x,z) - P_{\text{ref}}(x,z)|}{\max_{x,z}|P_{\text{ref}}(x,z)|}. \tag{10}$$

Thus, plots of $\eta(x,z)$ show the spatial distribution of error values, and $\eta_{\max}$ condenses all of the errors for each pressure field calculation into a single value.

IV. RESULTS

An example of a simulated beam pattern produced by a rectangular piston is presented in Fig. 2. This rectangular piston, which has a width of $2a=5\lambda$ and a height of $2b=7.5\lambda$, is equivalent to that simulated in Fig. 13 of Lockwood and Willette[2] with an aspect ratio $b/a=1.5$ and half-width $a=2.5\lambda$. In Fig. 2, the acoustic field is evaluated in the half-plane defined by $y=0$ and $x\geqslant 0$, which is located across the middle of the rectangular piston in the height direction and halfway across the face of the radiating piston in the width direction.

The half-plane in Fig. 2 is obtained when Eq. (2) is superposed with the appropriate choice of signs in Eq. (3). In this figure, Eq. (2) is evaluated with 200 000-point Gauss

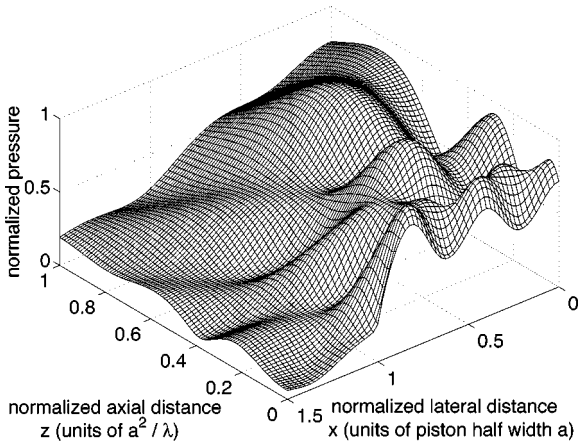


FIG. 2. Simulated time-harmonic pressure field for a 5λ wide and 7.5λ high rectangular source. The wavelength of the acoustic excitation is represented by λ, so the element half width is 2.5λ, and the aspect ratio is $b/a=1.5$. This simulated pressure distribution, which serves as the reference for subsequent error calculations, is obtained when Eq. (2) is evaluated with 200 000-point Gauss quadrature for each of the subelements and then superposed according to Eq. (3).

quadrature. This result is demonstrated as a reference because the maximum normalized error between this beam pattern and other beam patterns that are computed with a greater number of Gauss abscissas is $\eta_{\max}\approx 10^{-15}$. Larger errors are encountered if the Fourier transform of the impulse response in Eq. (2) is obtained using a substantially smaller number of Gauss abscissas. When the 200 000-point Gauss quadrature results using Eq. (2) are compared with 200 000-point Gauss quadrature applied to the rapid integral expression in Eq. (6), the mesh plots are indistinguishable, and the value of $\eta_{\max}$ for each is $\approx 10^{-15}$. In other words, the impulse response in Eq. (2) and the analytically equivalent formulation in Eq. (6) have converged to the same value at all points in space when each integral is evaluated with 200 000-point Gauss quadrature.

A. Spatial error distribution

Figure 3 shows the normalized difference $\eta(x,z)$ between the reference beam pattern in Fig. 2, which was computed with 200 000-point Gauss quadrature, and the Fourier transform of the impulse response described by Eq. (2) evaluated with 9-point Gauss quadrature. In Fig. 3, the spatial error distribution $\eta(x,z)$ consists of a peak value near the edge coordinates of the rectangular source $(x,z)=(a,0)$ and a numerical artifact that begins near the location of the peak error value and continues out in the $+z$ direction indefinitely along the line $x/a=1$. The artifact is barely noticeable on either side of the line $x/a=1$ in the nearfield region for calculations using fewer Gauss abscissas because the errors that the impulse response encounters elsewhere in the field are initially very large. Once the integrand is sufficiently sampled, the edge artifact emerges, and for the 5λ $\times$7.5λ rectangular source geometry considered here, the edge artifact becomes evident in this nearfield grid only when the impulse response is evaluated with nine or more Gauss abscissas.

Figure 4 contains the normalized difference between the reference beam pattern in Fig. 2 and the results of 9-point

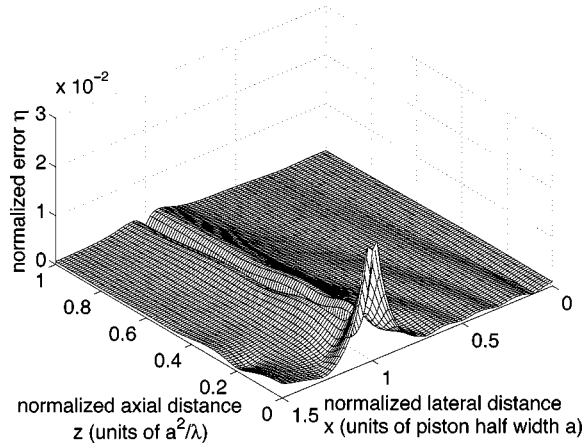


FIG. 3. Normalized difference between the reference beam pattern in Fig. 2 and the results of 9-point Gauss quadrature applied to the impulse response calculation in Eq. (2). A peak error value of 2.4×10^{-2} is indicated at the edge of the rectangular source. A numerical artifact, located on both sides of the line $x/a=1$, also extends across the mesh. If the number of Gauss abscissas is increased, the peak error in the neighborhood of $z=0$ disappears, but the remaining errors on either side of the line $x/a=1$ are only diminished slightly. The remaining errors that start near each edge of the source and continue outward in the $+z$ direction beyond the far edge of the mesh are produced by the inverse cosine terms in the impulse response.

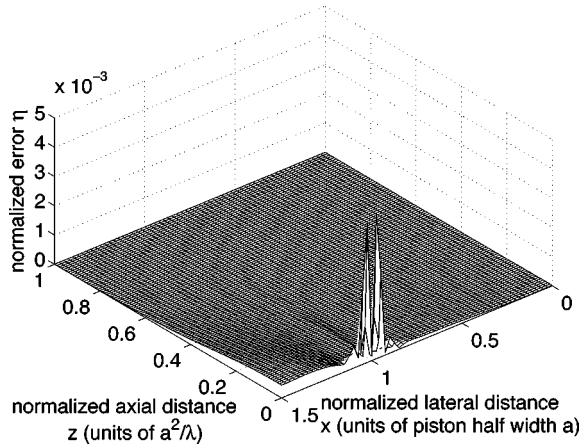


FIG. 4. Normalized difference between the reference beam pattern in Fig. 2 and the results of 9-point Gauss quadrature applied to the rapid formulation in Eq. (6). At the edge of the rectangular source, the peak error value reaches 5.0×10^{-3}, which is smaller than the error computed with the same number of Gauss abscissas demonstrated in Fig. 3 for the impulse response. Furthermore, the numerical artifact on both sides of the line $x/a=1$ in Fig. 3 is eliminated by subtracting the singularity in Eq. (6). As the number of Gauss abscissas is increased, the computed error values throughout the field are rapidly reduced.

Gauss quadrature applied to the rapid integral expression in Eq. (6). Throughout the computed field, the error values in Fig. 4 are smaller than those encountered Fig. 3, suggesting that the rapid formulation converges more quickly than the impulse response. Figure 2 also shows that the rapid approach eliminates the singularity in the impulse response that was demonstrated in Fig. 3 along the line $x/a=1$. Although some numerical errors remain in Fig. 4 near $z=0$, these are confined to a small region next to the face of the rectangular source. The spatial error distribution in Fig. 4 is representative of that obtained with four or more Gauss abscissas. As the number of Gauss abscissas increases, the error rapidly decreases everywhere in the pressure grid.

B. Error values and computation times

Figures 5 and 6 contain a summary of the errors calculated and times measured when results using the impulse response in Eq. (2) and the rapid formulation in Eq. (6) are evaluated numerically. In Figs. 5 and 6, the errors and run times are evaluated as the number of Gauss abscissas increases from 1 to 200. The computations are performed on an 800 MHz Intel Pentium III personal computer running the Red Hat Linux operating system version 7.1. On this computer, all simulations are run sequentially with the intent of maintaining similar load conditions for outside processes. Each simulation is written in the C programming language. All simulation routines, including those involving Eq. (2), apply every available numerical acceleration technique.

Figure 5 demonstrates that the peak errors obtained from the rapid integral expression in Eq. (6) are consistently smaller than those computed using the impulse response in Eq. (2). In Fig. 5, the reference beam is again defined in Fig. 2 as the result of 200 000-point Gauss quadrature applied to Eq. (2) and the maximum error is computed with Eq. (10). Figure 5 shows that the impulse response converges slowly at first, then accelerates over a short range until the singularity near the line $x/a=1$ is encountered. The impulse response again converges slowly once the singularity is observed. Figure 5 also demonstrates that the rapid formulation in Eq. (6) achieves a consistent reduction in the error as the number of Gauss abscissas increases and that no singularity is present along the line $x/a=1$.

Figure 6 shows that the computation time is linearly proportional to the number of Gauss abscissas. Figure 6 indicates that, for a 61 by 101 point rectilinear grid layout and a 5λ wide by 7.5λ wide rectangular source, the computation time required for calculations using the rapid integral formulation in Eq. (6) is only half that required for the impulse response. This comparison, which is evaluated independent of the respective maximum error values, only considers the time required to evaluate each expression for a certain number of Gauss abscissas applied to this grid and source geometry.

If these computation time comparisons are normalized with respect to the peak error values, the rapid formulation demonstrates even better performance. For example, the rapid integral method of Eq. (6) first achieves a peak error value below 10% with six Gauss abscissas, and the impulse response in Eq. (2) first reaches a peak error less than 10% with nine Gauss abscissas. After the computation time for the rapid method using six Gauss abscissas is divided into the computation time for the impulse response using nine Gauss abscissas, the result indicates that the rapid approach is 2.6 times faster than the impulse response for an error value target of 10%. Likewise, the peak error calculated for the impulse response drops below 1% with ten Gauss abscissas, and the peak error for the rapid method falls below the same peak error target with eight Gauss abscissas. This yields a ratio of computation times equal to 2.6, so the rapid approach is 2.6 times faster for a desired peak error of 1%. A peak error value of 0.1% is reached when 25 Gauss abscissas

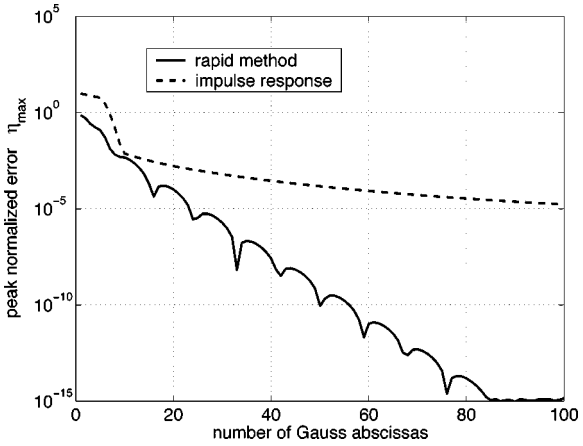


FIG. 5. Comparison of the maximum normalized errors η_{max} obtained from the rapid approach (solid line) and the impulse response (dashed line) as a function of the number of Gauss abscissas. This figure shows that, for the source and grid geometry evaluated here, the maximum error η_{max} computed with the rapid approach is consistently smaller than that obtained with the impulse response. With the rapid formulation in Eq. (6), the computed errors at individual grid locations are also typically smaller than those calculated with the impulse response.

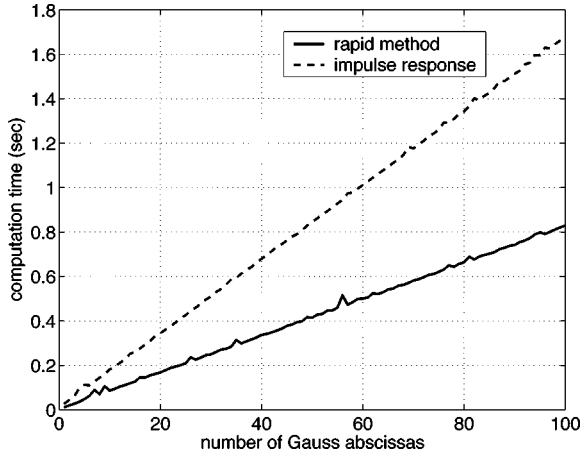


FIG. 6. Comparison of measured computation times required for simulations of a 5λ×7.5λ rectangular source using results obtained from the impulse response in Eq. (2), indicated with a dashed line, and the rapid formulation Eq. (6), indicated with a solid line, as a function of the number of Gauss abscissas. This result, along with the result of Fig. 5, shows that the rapid formulation in Eq. (6) simultaneously reduces the errors and the run times relative to those obtained with the impulse response.

are applied to the impulse response, and this same error is achieved when only 14 Gauss abscissas are applied to the rapid method. The quotient of the two run times is again evaluated, and the result indicates that the rapid approach is 3.6 times faster for a peak error of 0.1%. The computation times and peak errors are functions of the source and grid geometry, so some variations in these values are expected as the input parameters change.

V. DISCUSSION

A. Computation times

Although the overall structure of the routines tested for Eqs. (2) and (6) was generally the same, certain features of the integrals in Eq. (6) are responsible for the shorter run times achieved by the rapid formulation for the same number of Gauss abscissas. This result is demonstrated in Fig. 6. The impulse response Eq. (2) contains an inverse trigonometric function, and evaluating the inverse cosine term in Eq. (2) requires more computation time than division by a second order polynomial in Eq. (6). Other inverse trigonometric functions such as inverse sine and inverse tangent are also slower than division by a second-order polynomial. If any of these inverse trigonometric functions are applied to pressure calculations using the impulse response approach, then the rapid method is consistently faster, even if the pressure is only evaluated at a single point.

After each integral sharing an appropriate common integrand is combined for Eqs. (2) and (6), a reduction in the computation time is achieved for both methods. Additional reductions in the computation time are realized for the rapid method when the shared terms in Eq. (6) are exploited. The expressions σ^2+s^2 and σ^2+l^2 appear twice in Eq. (6), but each is only evaluated once and then stored in memory for subsequent calculations. The values of s and l remain unchanged as the computational grid is traversed in the $+z$ direction for a constant value of (x,y), and this allows the repeated use of the stored values for σ^2+s^2 and σ^2+l^2. The value of e^{-jkz} is also computed once for each value of z in the computational grid and then stored. When the corresponding z value is encountered in the evaluation of Eq. (6), the stored value for e^{-jkz} is inserted. Individually, each step reduces the computation time somewhat, but by combining all of these steps, the rapid method achieves a significant overall time savings.

For both the impulse response and the rapid method, the Gauss weights and abscissas are calculated in advance and stored in a file. These are the only values that remain unchanged as the grid and source geometries are varied, and these are the only values that are common to both pressure field calculation methods. These values are loaded prior to each calculation, so the time required to calculate the Gauss parameters is not included in Fig. 6. However, the Gauss weights and abscissas are the only values that are calculated in advance. After the complete mesh of pressure field values is calculated for a certain number of Gauss abscissas, all values in memory are cleared, and the next calculation starts from scratch. Therefore, each time value shown in Fig. 6 includes the time required for a complete pressure field calculation along with all of the associated overhead. In other words, by clearing memory after each pressure field mesh is calculated, each time value shown in Fig. 6 for a certain number of Gauss abscissas represents the computation time that is required for a entire mesh of pressure field values calculated from start to finish.

Although the rapid method reduces the computation time by factors of 2.6, 2.6, and 3.6 relative to the impulse response for 10%, 1%, and 0.1% peak errors, respectively, this result is specific to the grid and source geometry shown in Figs. 1 and 2. Other grid and source geometries can either reduce or increase the relative computation time. A smaller decrease in the computation time is expected for reduced grid sizes, whereas the rapid method is significantly faster when applied to larger three-dimensional (3D) grids. A larger

grid enables more shared calculations, and this in turn reduces the computation time required for the rapid method.

Computations of pressure fields on very large 3D grids are commonly performed in simulation studies of ultrasound phased arrays designed for thermal therapy. These ultrasound phased arrays consist of hundreds or thousands of independent elements, and the computed 3D pressure grid can reach hundreds of wavelengths on each side. Computations of the pressure fields generated by ultrasound therapy arrays include a substantial nearfield component. These calculations are also very time-intensive. Simulations of thermal therapy arrays and other large phased array systems will benefit greatly from the fast computation times and small numerical errors achieved by the rapid pressure field calculation method in Eq. (6).

B. Numerical errors

The numerical errors in Figs. 3–5 are caused by the aliasing or undersampling of each integrand. The primary source of aliasing in pressure field calculations using the impulse response is the inverse cosine function in Eq. (2). The inverse cosine term produces slopes that are particularly large where $x/a \approx 1$ and/or $y/b \approx 1$, and, in these regions, finite sampling rates amplify aliasing problems caused by the impulse response. The inverse tangent function encounters the same aliasing problems in these regions. In fact, the errors shown in Fig. 3 are produced by every pressure field calculation formula that uses inverse trigonometric functions. Along the line $x/a=1$ in the $y=0$ plane, the numerical errors are particularly severe in Fig. 3. As indicated by the result in Fig. 3, large slopes even cause numerical problems for Gauss quadrature, which provides additional samples near the upper and lower limits of integration. In the two-dimensional result shown in Fig. 3, the numerical errors produced by the impulse response are concentrated along a line tangent to the edge of the rectangular source, and in three dimensions, these errors are adjacent to the planes tangent to the edge of the source. The tangents to the four edges of the rectangular source are defined by $x=\pm a$ and $y=\pm b$ in the coordinate system described in Fig. 1, so the errors produced by the impulse response are found immediately adjacent to these four planes. In two-dimensional pressure fields, the errors are observed along lines as demonstrated in Fig. 3.

Equation (6) eliminates these errors by subtracting a term from the numerator of each integrand, therefore avoiding problems that are otherwise encountered when each denominator grows small. Thus, the singularity disappears after the e^{-jkz} term is subtracted in Eq. (6). The numerical error in Eq. (6) is generally dominated by the frequency components in the complex exponential term $e^{-jk\sqrt{z^2+\sigma^2+s^2}}$ in the integrand. If this complex exponential term is adequately sampled, then aliasing problems are avoided, and the numerical solution to Eq. (6) converges quickly as demonstrated in Fig. 5. A similar complex exponential term $e^{-jk\beta}$ in Eq. (2) also influences the convergence of the impulse response. If the $e^{-jk\beta}$ term is adequately sampled, aliasing problems are reduced in most locations away from the line at $x/a=1$ (or in 3D calculations, the planes $x=\pm a$ and $y=\pm b$). Near $x/a=1$, the impulse response requires a substantial number of additional samples to reduce the effects of aliasing introduced by the $\cos^{-1}$ term in Eq. (2).

These results suggest that the rapid method is better suited for calculations of the time-harmonic pressure field produced by a rectangular source than the impulse response. Results obtained with Eq. (2) occasionally achieve a smaller maximum error for the same number of Gauss abscissas in a limited number of locations for some restricted combinations of source and grid geometries; however, Eq. (6) produces significantly smaller errors in the vast majority of acoustic field calculations. Furthermore, computation times are consistently shorter with the rapid approach. Whenever the measured computation times are compared for the same maximum error, the rapid method is consistently faster than the impulse response.

C. Circular sources

Similar solutions for the time-harmonic fields generated by a circular piston are presented by Archer-Hall *et al.*[10] and Hutchins *et al.*[11] The solution in Ref. 10 applies a cylindrical coordinate system with a movable origin[1] to the solution of the Kirchhoff integral. These manipulations convert a double integral into a simplified single integral, and the resulting expression is similar to Eq. (6). The solution presented by Hutchins *et al.*[10] begins by deriving the velocity potential for a circular piston driven by a time-harmonic excitation. The velocity potential is expressed as a double integral containing Bessel functions, and this result is simplified using Hankel transform tables. When the best features of the Archer-Hall *et al.*[10] and Hutchins *et al.*[11] solutions are combined,[12] the resulting single integral expression is similar to Eq. (6).

D. Other integration techniques

Gauss quadrature,[9] when applied to the integrals presented herein, consistently produces more accurate results in less time than other standard integration procedures, including the trapezoidal rule, Simpson's rule, and Newton–Cotes rules. These integration techniques were tested extensively on Eqs. (2) and (6) for a fixed number of abscissas, and Gauss quadrature was retained after superior numerical accuracy was clearly established for these two integral formulations. Similar tests were also performed with an adaptive Romberg integration scheme, and the adaptive scheme failed to consistently converge within the specified tolerance, especially when applied to Eq. (2). Thus, Gauss quadrature is the preferred numerical integration scheme for Eqs. (2) and (6).

E. Future work

Further numerical improvements are still possible for time-harmonic simulations of the nearfield pressure produced by a rectangular piston. Future studies will include considerations of spatial sampling,[13–15] which changes the number of abscissas required as a function of the piston geometry and the field coordinates. If the number of Gauss abscissas is allowed to vary spatially in a scheme that maintains sufficient sampling rates for each integrand, then the computation time will be reduced while the computed pres-

sure fields maintain a specified maximum error. This is achieved with a grid sectoring scheme for circular pistons,[12] and similar results are expected for rectangular sources.

VI. CONCLUSION

Numerical singularities in the impulse response were identified and eliminated for a time-harmonic excitation applied to a rectangular source. These singularities, which emanate from the edges of a rectangular source, result from the large slopes produced by the inverse trigonometric functions in the impulse response. The numerical problems caused by these singularities are removed by substituting equivalent integral expressions and exchanging the order of integration in each. Further improvements are achieved by subtracting the singularity from each integrand, by exploiting repeated expressions, and by combining integrals which share integrands and a single common limit. The resulting expressions are evaluated with Gauss quadrature. The results show that the rapid approach in Eq. (6), when evaluated numerically, consistently outperforms the impulse response in Eq. (2) in terms of both computation time and numerical error. In calculations of the pressure field produced by the $5\lambda \times 7.5\lambda$ rectangular source shown in Fig. 2, the rapid approach is 2.6 times faster than the impulse response for maximum specified errors of 10% and 1%. Even greater reductions in the computation time are achieved by the rapid approach relative to the impulse response as the maximum specified error is reduced and as the size of the computational grid is increased.

ACKNOWLEDGMENTS

This work was funded in part by a grant from the Whitaker Foundation, NIH Grant No. 5PO1 CA42745, and NIH Grant No. 1R01 CA093669. The author would like to thank Dr. Thaddeus V. Samulski and Jeremy D. Hoff of the Department of Radiation Oncology at Duke University Medical Center for their helpful comments and suggestions.

[1] P. R. Stepanishen, ''Transient radiation from pistons in an infinite planar baffle,'' J. Acoust. Soc. Am. **49**, 1629–1638 (1971).
[2] J. C. Lockwood and J. G. Willette, ''High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled piston,'' J. Acoust. Soc. Am. **53**, 735–741 (1973).
[3] J. L. San Emeterio and L. G. Ullate, ''Diffraction impulse-response of rectangular transducers,'' J. Acoust. Soc. Am. **92**, 651–662 (1992).
[4] M. Arditi, F. S. Foster, and J. W. Hunt, ''Transient fields of concave annular arrays,'' Ultrason. Imaging **3**(1), 37–61 (1981).
[5] G. R. Harris, ''Transient field of a baffled planar piston having an arbitrary vibration amplitude distribution,'' J. Acoust. Soc. Am. **70**, 186–204 (1981).
[6] W. A. Verhoef, M. J. T. M. Cloostermans, and J. M. Thijssen, ''The impulse-response of a focused source with an arbitrary axisymmetric surface velocity distribution,'' J. Acoust. Soc. Am. **75**, 1716–1721 (1984).
[7] J. A. Jensen, ''A new calculation procedure for spatial impulse responses in ultrasound,'' J. Acoust. Soc. Am. **105**, 3266–3274 (1999).
[8] P. J. Davis and P. Rabinowitz, *Numerical Integration* (Academic, New York, 1975), pp. 73–76, 87–90, 139–140, 369.
[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1972), pp. 887–889, 916–919.
[10] J. A. Archer-Hall, A. I. Bashter, and A. J. Hazelwood, ''Means for computing the Kirchhoff surface integral for a disk radiator as a single integral with fixed limits,'' J. Acoust. Soc. Am. **65**, 1568–1570 (1979).
[11] D. A. Hutchins, H. D. Mair, P. A. Puhach, and A. J. Osei, ''Continuous-wave pressure fields of ultrasonic transducers,'' J. Acoust. Soc. Am. **80**, 1–12 (1986).
[12] R. J. McGough, T. V. Samulski, and J. F. Kelly, ''An efficient grid sectoring method for calculations of the nearfield pressure generated by a circular piston,'' J. Acoust. Soc. Am. (accepted, 2004).
[13] K. B. Ocheltree and L. A. Frizzell, ''Sound field calculation for rectangular sources,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36**(2), 242–248 (1989).
[14] D. P. Orofino and P. C. Pedersen, ''Multirate digital signal-processing algorithm to calculate complex acoustic pressure fields,'' J. Acoust. Soc. Am. **92**, 563–582 (1992).
[15] Z. G. Hah and K. M. Sung, ''Effect of spatial sampling in the calculation of ultrasonic fields generated by piston radiators,'' J. Acoust. Soc. Am. **92**, 3403–3408 (1992).

An efficient grid sectoring method for calculations of the near-field pressure generated by a circular piston

Robert J. McGough[a)]
Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan 48824

Thaddeus V. Samulski
Department of Radiation Oncology, Duke University Medical Center, Durham, North Carolina 27710

James F. Kelly
Department of Mathematics, Michigan State University, East Lansing, Michigan 48824

(Received 25 January 2004; accepted for publication 28 January 2004)

An analytical expression is derived for time-harmonic calculations of the near-field pressure produced by a circular piston. The near-field pressure is described by an efficient integral that eliminates redundant calculations and subtracts the singularity, which in turn reduces the computation time and the peak numerical error. The resulting single integral expression is then combined with an approach that divides the computational grid into sectors that are separated by straight lines. The integral is computed with Gauss quadrature in each sector, and the number of Gauss abscissas in each sector is determined by a linear mapping function that prevents large errors from occurring in the axial region. By dividing the near-field region into 10 sectors, the raw computation time is reduced by nearly a factor of 2 for each expression evaluated in this grid. The grid sectoring approach is most effective when the computation time is reduced without increasing the peak error, and this is consistently accomplished with the efficient integral formulation. Of the four single integral expressions evaluated with grid sectoring, the efficient formulation that eliminates redundant calculations and subtracts the singularity demonstrates the smallest computation time for a specified value of the maximum error. [DOI: 10.1121/1.1687835]

PACS numbers: 43.20.Rz, 43.20.El, 43.35.Bf [ADP] Pages: 1942–1954

I. INTRODUCTION

The impulse response of a circular piston[1–3] defines a linear input–output relationship of a single transducer with a closed-form single integral expression. This expression implicitly summarizes the temporal and spatial characteristics of a particular transducer geometry in a framework that is convenient for pulsed and time-harmonic acoustic field computations. These computations ordinarily assume that the velocity profile across the surface of the transducer is spatially uniform, although certain other models permit some radial variation in the velocity profile.[4,5] The impulse response is primarily used for pulsed mode calculations in the near-field region; however, this model is also suitable for time-harmonic pressure calculations in the near field.

Two other equivalent single integral expressions are described by Archer-Hall *et al.*[6] and Hutchins *et al.*[7] specifically for time-harmonic near-field calculations of the pressure generated by a circular piston. The expression in Archer-Hall *et al.*[6] applies a cylindrical coordinate system with a movable origin[2] to the solution of the Kirchhoff integral. These geometric manipulations convert a double integral into a simplified single integral. A similar expression derived by Hutchins *et al.*[7] computes the velocity potential for a circular piston driven by a sinusoidal excitation. The velocity potential is expressed as a double integral containing Bessel functions, and this result is simplified using Hankel transform tables. The resulting single integral expression[7] is closely related to the result of Archer-Hall *et al.*[6] Although these single integral expressions are all analytically equivalent, the numerical properties of these expressions are quite different. The single integral expressions of Archer-Hall *et al.*[6] and Hutchins *et al.*[7] are particularly amenable to certain numerical improvements.

After three single integral expressions for time-harmonic near-field pressures produced by a circular piston are briefly reviewed, an efficient single integral expression is derived from the results of Archer-Hall *et al.*[6] and Hutchins *et al.*[7] A sector-based technique for spatially varying the number of Gauss abscissas applied to each integral is also introduced. Results of near-field beam patterns produced by each expression are then compared for a fixed number of Gauss abscissas. All four expressions are evaluated and compared with a grid sectoring technique that varies the number of Gauss abscissas as the spatial grid coordinates change. The results show that the peak errors increase significantly when the impulse response and the expression derived in Archer-Hall *et al.*[6] are evaluated with grid sectoring. The peak errors increase over a limited range of values if the single integral expression of Hutchins *et al.*[7] is computed with grid sectoring. These numerical problems are avoided altogether in near-field calculations that combine the efficient expression

a)Electronic mail: mcgough@egr.msu.edu

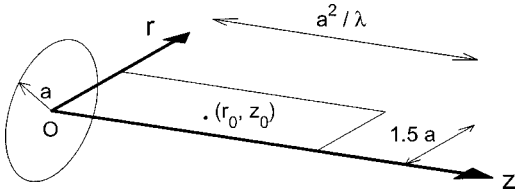


FIG. 1. Definition of coordinate axes for acoustic field calculations. The origin (O) of the coordinate system is defined as the intersection between the z and r axes in a cylindrical coordinate system. The z axis is coincident with the piston normal, and individual field coordinates are indicated by the pair (r_0,z_0). The radius of the circular piston is designated by a, the extent of the field calculation in the radial direction is $1.5a$, and the extent of the computed field in the axial direction is a^2/λ. These are the same coordinates used by Zemanek (Ref. 8) and Lockwood and Willette (Ref. 3) for calculations of pressure fields generated by circular pistons.

and grid sectoring. The efficient formulation, which combines the best features of the Archer-Hall *et al.*[6] and Hutchins *et al.*[7] expressions, achieves the smallest peak errors in the shortest time for calculations in a single sector. For near-field calculations evaluated in multiple sectors, the efficient single integral expression also demonstrates the smallest computation time for a specified value of the peak error.

II. SINGLE INTEGRAL EXPRESSIONS

The impulse response for a circular piston,[1–3] defined here as $h(r,z;t)$, is a function of temporal and spatial parameters. The impulse response of a circular piston is also applicable to time-harmonic pressure calculations once the Fourier transform, $H(r,z;k)=\mathcal{F}\{h(r,z;t)\}$, is evaluated. The Fourier transform of the impulse response is related to the steady-state pressure p generated by a circular piston according to

$$p(r,z;k)=-j\omega\rho v e^{j\omega t}H(r,z;k), \tag{1}$$

where ω is the excitation frequency in radians per second, ρ is the density of the medium, v is a constant normal velocity evaluated at the surface of the circular piston, k is the wave number, and t is the elapsed time. As indicated in Eq. (1), $H(r,z;k)$ is directly proportional to the pressure distribution, and all of the spatial variations in the pressure field are represented by the expression $H(r,z;k)$. Therefore, the quantity $H(r,z;k)$ is computed in the analysis that follows. The relationship between each computed pressure field and the circular piston is shown in Fig. 1, where the center of the circular piston is the origin of the cylindrical coordinates (r,z), and the positive z direction is coincident with the normal at the center of the circular piston. In Eq. (1), the Fourier transform of the impulse response H is represented as a function of the wave number $k=\omega/c$, which is implicitly related to the excitation frequency $f=\omega/(2\pi)$. This notation simplifies the following expressions for H, which is computed for a circular piston with several analytically equivalent but numerically dissimilar single integral expressions.

A. Oberhettinger,[1] Stepanishen,[2] and Lockwood and Willette[3]

In pressure field calculations for pulsed excitations, the impulse response is convolved with a temporally dependent excitation function, and therefore the corresponding convolution integral is necessarily evaluated with respect to the time variable. For a sinusoidal excitations, however, distance is a more convenient variable. Therefore, the notation describing the time limits in Lockwood and Willette[3] is adapted to distance limits here. The distance values required for time harmonic pressure field calculations using the impulse response are

$$\begin{aligned}\beta_1&=z,\\ \beta_2&=\sqrt{z^2+(r-a)^2},\\ \beta_3&=\sqrt{z^2+(r+a)^2}.\end{aligned} \tag{2}$$

These limits are then applied to the Fourier transform of the impulse response,

$$H_1(r,z;k)=\begin{cases}(e^{-jkz}-e^{-jk\sqrt{z^2+a^2}})/(jk), & r=0\\ I_1+(e^{-jkz}-e^{-jk\beta_2})/(jk), & 0<r<a,\\ I_1, & r\geqslant a\end{cases} \tag{3}$$

where the integral I_1 is defined as

$$I_1=\frac{1}{\pi}\int_{\beta_2}^{\beta_3}\cos^{-1}\left\{\frac{\beta^2-z^2+r^2-a^2}{2r\sqrt{\beta^2-z^2}}\right\}e^{-jk\beta}d\beta. \tag{4}$$

These expressions, which describe the pressure field produced by a sinusoidal excitation, are valid at all grid coordinates where zzo. Equations (3) and (4) are obtained from the impulse response presented in Oberhettinger,[1] Stepanishen,[2] and Lockwood and Willette[3] after the change of variable $t=\beta/c$ is applied, where c represents the speed of sound.

B. Archer-Hall *et al.*[6]

An equivalent expression for Eq. (3), which applies the coordinate geometry suggested by Stepanishen[2] to the integral of the time-harmonic Green's function evaluated over the face of a circular piston, is derived in Archer-Hall *et al.*[6] This expression is presented in condensed form[7] as

$$H_2(r,z;k)=\frac{a}{jk\pi}I_2+\frac{1}{jk}e^{-jkz}\begin{cases}1, & r<a\\ \frac{1}{2}, & r=a\\ 0, & r>a,\end{cases} \tag{5}$$

where I_2 is the integral expression defined by

$$I_2=\int_0^{\pi}\frac{(r\cos\psi-a)}{r^2+a^2-2ar\cos\psi}e^{-jk\sqrt{r^2+a^2-2ar\cos\psi+z^2}}d\psi. \tag{6}$$

In Eq. (5), only the $(1/jk)e^{-jkz}$ term is multiplied by 1, 1/2, or 0, depending on the location of the radial coordinate r relative to the edge of the piston. The integral I_2 is evaluated numerically in all off-axis locations. At $r=0$, the $\cos\psi$ contributions disappear from Eq. (6), and the resulting value of I_2 is $-(\pi/a)\exp(-jk\sqrt{a^2+z^2})$. This result is exactly equal to the corresponding on-axis term in Eq. (3).

C. Hutchins *et al.*[7]

Another analytically equivalent expression is derived from a Green's function approach[7] that generates a double integral containing Bessel functions. This result is simplified using Hankel transform tables, which produces

$$H_3(r,z;k)=\frac{a}{j2\pi k}\int_{-\pi}^{\pi}\frac{a-re^{-j\psi}}{a^2+r^2-2ar\cos\psi}\times(e^{-jkz}-e^{-jk\sqrt{r^2+a^2-2ar\cos\psi+z^2}})d\psi. \quad (7)$$

As noted in Hutchins *et al.*,[7] the ratio $(a-re^{-j\psi})/(a^2+r^2-2ar\cos\psi)$ in Eq. (7) is analytically equivalent to $(a-re^{j\psi})^{-1}$. Of these two choices, the former is preferred for numerical calculations in languages (such as C) that fail to provide explicit compiler support for complex arithmetic.

III. NUMERICAL TECHNIQUES

Several techniques improve the numerical performance of these single integral expressions for the Fourier transform of the impulse response evaluated for a circular piston. Some of these techniques are applicable to all of the formulations, and others are only applicable to specific expressions. Each integral is evaluated with Gauss quadrature, and normalized errors are computed for each result.

A. Efficient formulation

Although H_2 and H_3 are analytically equivalent, each expression possesses unique numerical properties. By combining the best numerical features of H_2 and H_3, an efficient single integral formulation with improved numerical performance is obtained. The derivation of the efficient expression begins either with H_2 or H_3, and each converges to the same result, yielding a fourth equivalent single integral expression for the Fourier transform of the impulse response.

One derivation begins with the expressions for H_2 and I_2 in Eqs. (5) and (6), respectively. The integral I_2 contains a numerical singularity that is encountered whenever $r\approx a$ and $\psi\approx 0$, and this singularity is responsible for increasing the numerical errors in all spatial locations where $r\approx a$. The numerical errors increase whenever a very small number in the denominator of the integrand in I_2 is divided into a much larger number in the numerator. Although division by zero is avoided either with a simplified expression for the integrand or with the selection of a numerical integration method that excludes the end points of the interval (such as Gauss quadrature), increased numerical errors are guaranteed with the Archer-Hall *et al.*[6] expressions for H_2 and I_2 in all locations where $r\approx a$. The numerical problems are eliminated by subtracting a singularity[9] from the exponential term in the numerator of I_2. The required term e^{-jkz} is obtained from a series expansion and then subtracted from the complex exponential term $e^{-jk\sqrt{r^2+a^2-2ar\cos\psi+z^2}}$ in I_2. The resulting expression, which instead divides a small number by a small number when $r\approx a$ and $\psi\approx 0$ and therefore reduces the error where $r\approx a$, is retained for numerical calculations. This same term is also added to the complex exponential term in I_2 and evaluated analytically. The analytical result exactly cancels the term on the far right-hand side of H_2 in Eq. (5), yielding the efficient formulation

$$H_4(r,z;k)=\frac{a}{jk\pi}\int_0^{\pi}\frac{r\cos\psi-a}{r^2+a^2-2ar\cos\psi}\times(e^{-jk\sqrt{r^2+a^2-2ar\cos\psi+z^2}}-e^{-jkz})d\psi. \quad (8)$$

The same result is also obtained from H_3 through manipulations which are very similar to those presented by Hutchins *et al.*[7] in a derivation that equates H_3 and H_2. That derivation eliminates the e^{-jkz} term from the integrand; however, this term is retained here because subtracting the singularity reduces the error in all locations where $r\approx a$. The first step in this derivation notes that the real part of the ratio $\gamma=(a-re^{-j\psi})/(r^2+a^2-2ar\cos\psi)$ in H_3 is an even function. Likewise, the imaginary part of γ is an odd function with respect to ψ for all values of a and r. Furthermore, both the real and the imaginary parts of the difference $(e^{-jk\sqrt{r^2+a^2-2ar\cos\psi+z^2}}-e^{-jkz})$ in H_3 are even functions with respect to ψ for all values of a, r, and z. The integral with respect to ψ of the imaginary part of γ (an odd function) multiplied by the difference between the two complex exponential functions (an even function), evaluated from $-\pi$ to π, is equal to zero. Thus, the imaginary part of γ is eliminated. The second step notes that the integral with respect to ψ of the real part of γ (an even function) multiplied by the difference between the two complex exponential functions (an even function), evaluated from $-\pi$ to π, is equal to twice the integral of the same expression from 0 to π. The result of these manipulations again yields the efficient expression H_4 in Eq. (8). Although the expressions H_2, H_3, and H_4 are similar in appearance, the numerical performance of each expression is quite different, especially when grid sectoring schemes are defined for numerical integration.

B. Grid sectoring

In each of the expressions H_1, H_2, H_3, and H_4, the spatial distribution of the numerical error is nonuniform when the integrand is evaluated with a constant number of abscissas in all field locations. This suggests that the integrand is oversampled in some regions and undersampled in others. Increasing the number of abscissas in regions where the error is large and decreasing the number of abscissas in regions where the error is small improves the efficiency of near-field pressure calculations. This result is ordinarily accompanied by a reduction in the computation time.

Grid sectoring, when applied to certain integral formulations, reduces the time required for pressure field calculations without increasing the maximum numerical error. This approach divides the computational grid in Fig. 1 into sectors, and the number of abscissas utilized in the numerical evaluation of the integral expression for $H(r,z;k)$ is determined by the sector in which a particular grid point is located. Grid sectoring only maintains the maximum error value for integral expressions that avoid numerical singularities.

A similar multirate sampling scheme in Orofino and Pedersen[10] defines regions based on contours that are extracted from the argument of the exponential term in I_1. This scheme, expressed here in terms of the parameters required for time-harmonic calculations using the impulse response, notes that the argument of $e^{-jk\beta}$ in Eq. (4), evaluated from β_2 to β_3, ranges from $k\sqrt{(r-a)^2+z^2}$ to $k\sqrt{(r+a)^2+z^2}$. The extent of this range is then $k\sqrt{(r+a)^2+z^2} - k\sqrt{(r-a)^2+z^2}$, and a proportional number of abscissas N is specified according to

$$N = \frac{C}{\lambda}\left(\sqrt{(r+a)^2+z^2} - \sqrt{(r-a)^2+z^2}\right), \tag{9}$$

where λ is the wavelength and C is a scale factor.

The grid sectoring approach employed here improves the multirate scheme of Orofino and Pedersen[10] by replacing curved contour boundaries with sectors defined by straight lines. The expression for the linear sector boundaries is derived from

$$N \approx \frac{C2ar}{\lambda\sqrt{r^2+z^2}}, \tag{10}$$

which is an asymptotic approximation to Eq. (9). Equation (10) is obtained after a binomial expansion is evaluated for both square root terms in Eq. (9). The approximation $r^2+z^2 \gg a^2$ is applied to the result. In Eq. (10), the spatially varying term $r/\sqrt{r^2+z^2}$, is exactly equal to $\sin\theta$ in the coordinate system of Fig. 1. Constant values of θ, which are indicated by straight lines emanating from the center of the circular piston, and are measured with respect to the piston normal define sector boundaries between successive values of N. Incremental changes in N are therefore related to θ through the $\sin^{-1}$ function. Likewise, $\tan\theta = r/z$ in Fig. 1, so the boundaries between adjacent sectors are expressed by

$$r/z = \tan\left(\sin^{-1}\left\{\frac{i-1}{n_l-1}\right\}\right) \quad \text{for } i=1 \text{ to } n_l. \tag{11}$$

Equation (11) evaluates the argument of $\sin^{-1}$ at equal increments of the range of (r,z) pairs considered, and the boundaries thus defined are a series of n_l lines which divide the computational grid in Fig. 1 into $n_s = n_l - 1$ sectors.

Figure 2 contains a comparison of the contours indicated by Eq. (9) and the sectors specified by Eq. (11). The solid lines are evaluated according to Eq. (11) for $n_s = 10$ sectors. In Fig. 2, two of the $n_l = 11$ solid lines are coincident with the r and z axes, and therefore only 9 of the solid lines are apparent. The solid lines are asymptotically related to the 9 contours indicated by the dashed lines in Fig. 2. These contours, which are computed directly from Eq. (9), are equally spaced over the range of values generated by Eq. (9). In Fig. 2, the solid lines from Eq. (11) are consistently closer to the z axis than the corresponding dashed contours from Eq. (9). This increases the size of the regions that require the most abscissas, which in turn reduces the errors near the boundaries between adjacent sectors. Also, Eq. (11) is more convenient than Eq. (9) for numerical calculations performed on a rectilinear grid. Intersections between grid lines and the lines defined in Eq. (11) are easily converted into limits defined for loop constructs, both for aligned and rotated grid orientations with respect to the piston source.

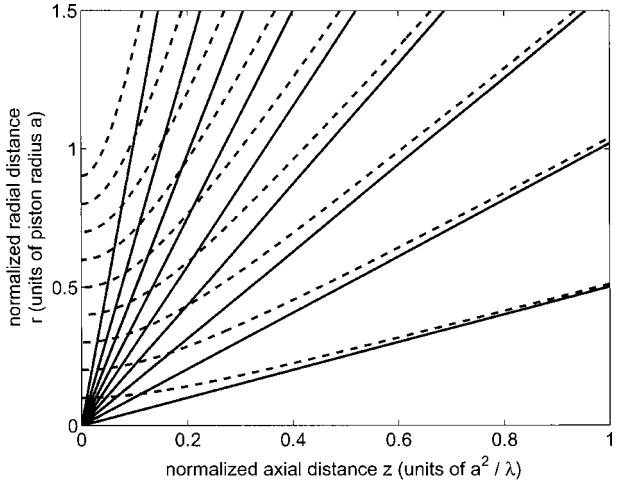


FIG. 2. Equation (11) predicts the asymptotic behavior of Eq. (9) for a circular piston with radius $a = 5\lambda$. Here, the solid lines define the asymptotes indicated by Eq. (11) for $n_l = 11$ straight lines dividing the computational grid into $n_s = 10$ separate sectors. Two of the lines are coincident with the z and r axes, respectively, so only 9 straight lines are evident. The asymptotes are superposed over a contour plot of Eq. (9) evaluated for the same piston radius. The dashed lines show where 9 equally spaced contours are located in the same computational grid. The solid lines obtained from Eq. (11), which are easier to compute than the dashed contours, define an efficient grid sectoring scheme for pressure field calculations.

The numerical error is also reduced by increasing the number of abscissas in the axial region. Equations (9) and (10) underestimate the number of abscissas required for near-field pressure calculations in the axial region, which motivates the construction of an expression that specifically increases the number of abscissas in the sectors nearest the piston normal. This is achieved with a linear mapping function

$$M = \frac{N_{\max}(1-S)}{n_s-1}(i-n_s) + N_{\max} \quad \text{for } i=1 \text{ to } n_s, \tag{12}$$

where $S \in [0,1]$ is a scale factor, $N_{\max}$ is the specified largest number of Gauss abscissas, n_s is the total number of sectors defined by Eq. (11), and M is the number of abscissas in the ith sector. With Eq. (12), the number of abscissas in each sector ranges from a minimum value of $N_{\max}S$ to a maximum value of $N_{\max}$. For a piston with radius $a = 5\lambda$, $S = 0.3$ is an appropriate value for calculations with $n_s = 10$ sectors. Likewise, the number of sectors n_s defined in the near-field region typically ranges from 5 to 15, depending on the value of S as well as the piston and grid geometry.

C. Gauss quadrature

The Gauss quadrature rule[11] evaluates the integrals in Eqs. (3), (5), (7), and (8). Gauss quadrature achieves lower errors than the trapezoidal rule, Simpson's rule, and other standard numerical integration schemes using the same number of abscissas. For results normalized to a specified maximum error value, Gauss quadrature is generally faster than other numerical integration methods, including adaptive methods (e.g., Romberg integration, etc.) which sometimes

fail to converge for impulse response calculations where $r \approx a$. The Gauss quadrature rule is defined according to the descriptions given by Davis and Rabinowitz[11] and Abromowitz and Stegun.[12] The Gauss rule computes the abscissas g_i and weights w_i of the N point Gauss–Legendre integration rule for the interval $[-1,1]$, and the abscissas are then converted by a linear mapping function[12] defined for an arbitrary input interval $[\alpha_{\min},\alpha_{\max}]$ according to

$$\alpha_i = \frac{\alpha_{\max}-\alpha_{\min}}{2} g_i + \frac{\alpha_{\max}+\alpha_{\min}}{2}, \tag{13}$$

where the Gauss weights w_i are multiplied by the scalar quantity $(\alpha_{\max}-\alpha_{\min})/2$. In Eq. (13), the pair $[\alpha_{\min},\alpha_{\max}]$ represents $[\sqrt{z^2+(r-a)^2},\sqrt{z^2+(r+a)^2}]$ for evaluations of H_1 in Eq. (3), $[\alpha_{\min},\alpha_{\max}]$ represents $[0,\pi]$ for H_2 and H_4 in Eqs. (5) and (8), respectively, and $[\alpha_{\min},\alpha_{\max}]$ represents $[-\pi,\pi]$ for H_3 in Eq. (7).

D. Error calculations

The numerical error $\eta(r,z)$ in the computed acoustic field is defined here as the normalized absolute value of the difference between a specific beam pattern and a separate reference beam pattern. The spatial distribution of the error is represented by

$$\eta(r,z) = \frac{|H(r,z)-H_{\text{ref}}(r,z)|}{\max_{r,z}|H_{\text{ref}}(r,z)|} \tag{14}$$

and the maximum error is defined as

$$\eta_{\max} = \frac{\max_{r,z}|H(r,z)-H_{\text{ref}}(r,z)|}{\max_{r,z}|H_{\text{ref}}(r,z)|}. \tag{15}$$

In Eqs. (14) and (15), $H_{\text{ref}}(r,z)$ represents the reference beam pattern, and $H(r,z)$ describes the Fourier transform of the impulse response that is compared to the reference field. Although these expressions implicitly depend on the wave number k, the wave number is constant in all of the calculations that follow, so only the spatial coordinates are included. Here, $H(r,z)$ is interchangeable with the pressure distribution $p(r,z)$ since all results are normalized and the leading constant terms in Eq. (1) cancel in the numerator and denominator of Eqs. (14) and (15). The scalar normalization factor in both equations prevents division by zero and avoids exaggerating differences where the field amplitudes are relatively small. The resulting spatial variations in the error $\eta(r,z)$ are presented in mesh plots where peak values indicate the presence of possible numerical singularities.

IV. SIMULATION RESULTS

All simulations routines were written in the C programming language. The simulations were performed on a generic 800 MHz Pentium III personal computer running the Red Hat Linux 7.1 operating system. On this computer, all simulations were run sequentially, and outside processes were limited to maintain similar load conditions at all times. The compiler switches were set to a standard level of optimization for all expressions and then left alone. No other optimization settings were examined.

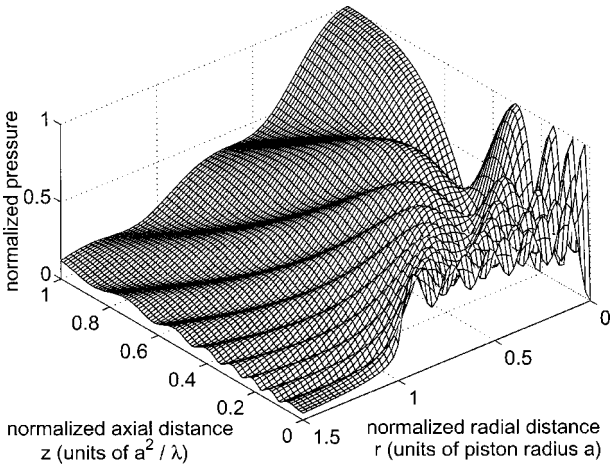


FIG. 3. Simulated acoustic field generated when 80 000-point Gauss quadrature is applied to the impulse response calculation described in Eq. (3) and represented by H_1. The wavelength of the acoustic excitation is equal to λ, and the radius of the circular piston is $a=5\lambda$. The spatial axes are normalized in the radial direction with respect to the piston radius a and in the axial direction with respect to a^2/λ. This simulation result is the reference for all subsequent error calculations in Figs. 4, 6, and 8–11.

A. Reference beam pattern

An example of a simulated beam pattern generated by a circular piston is presented in Fig. 3. The radius of the circular piston is $a=5\lambda$, which is the same value used in Fig. 8 of Lockwood and Willette,[3] Fig. 2(b) in Hutchins *et al.*,[7] and Fig. 6 in Zemanek.[8] In Fig. 3, the acoustic field is evaluated in the half-plane bounded by the line $r=0$ in one direction and the line $z=0$ in the other direction. This figure, which evaluates the impulse response expression H_1 at every field coordinate with 80 000-point Gauss quadrature, is the reference beam pattern for all error calculations. Larger maximum normalized differences are encountered if fewer Gauss abscissas are applied to H_1, whereas the maximum normalized difference never exceeds $\eta_{\max}=10^{-13}$ as the number of Gauss abscissas increases above $M=80\,000$. In addition, the results of the other three equations evaluated with 80 000-point Gauss quadrature are indistinguishable from Fig. 3. The expressions H_1, H_2, H_3, and H_4 generate solutions that are within at most $\eta_{\max}=10^{-13}$ of one another at all points in the simulated field presented in Fig. 3, which indicates that all four expressions converge to the same result if the pressure field is computed with 80 000 or more Gauss abscissas.

B. Errors and times for a single sector ($n_s=1$)

The convergence of H_1, H_2, H_3, and H_4 is established with 80 000 Gauss abscissas; however, acceptable time-harmonic calculations of the near-field pressure are obtained with a smaller number of Gauss abscissas. This is demonstrated in Fig. 4, which displays the peak error values $\eta_{\max}$ computed for a circular piston with $a=5\lambda$ using all four single integral expression expressions. For each value plotted in Fig. 4, the entire near-field pressure distribution is computed with the same number of Gauss abscissas at each grid coordinate, and peak error values are calculated with Eq. (15). In Fig. 4, near-field calculations are repeated for each of the expressions H_1, H_2, H_3, and H_4, and for each ex-

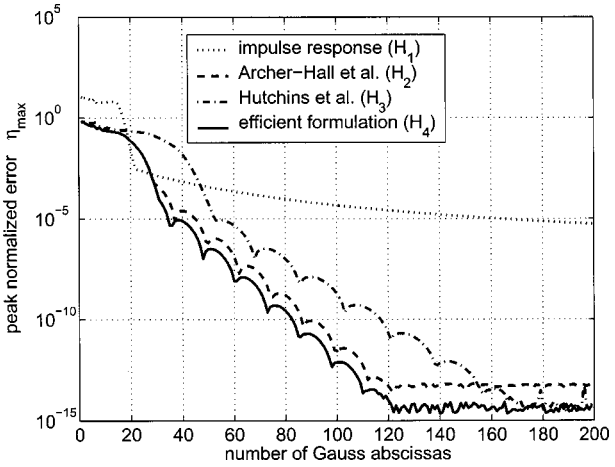


FIG. 4. A comparison of the peak normalized errors $\eta_{\max}$ obtained from beam patterns computed with H_1, H_2, H_3, and H_4 and evaluated for a constant number of Gauss abscissas in a single sector. Each beam pattern is computed for a circular source with radius $a=5\lambda$. The errors are computed relative to the reference in Fig. 3. The efficient formulation H_4 (solid line) generally produces the smallest peak errors $\eta_{\max}$, and the peak errors obtained with the Archer-Hall *et al.* (Ref. 6) formulation H_2 (dashed line) are very similar if the same number of Gauss abscissas are applied across the entire computational grid. Likewise, the peak errors obtained with the Hutchins *et al.* (Ref. 7) formulation H_3 (dot dash) are consistently larger than those obtained with H_2 and H_4. The errors obtained with the impulse response H_1 (dotted line) are similar to those achieved with the efficient method H_4 over a very small range of values, although for most values, the peak errors $\eta_{\max}$ are largest with the impulse response H_1.

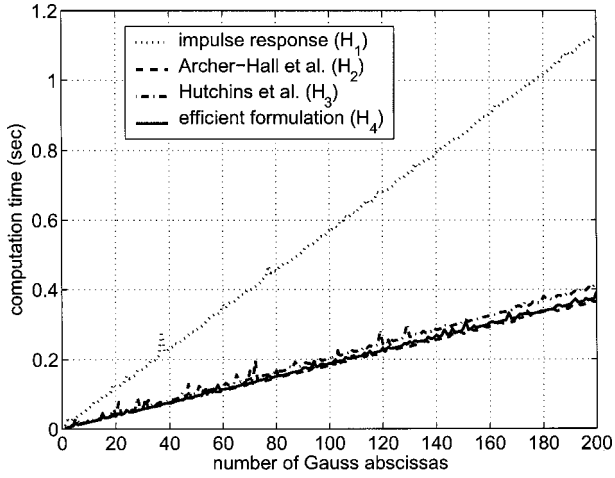


FIG. 5. A comparison of measured computation times for beam patterns calculated with a constant number of Gauss abscissas using the expressions H_1, H_2, H_3, and H_4. As in Fig. 4, each beam pattern is computed on the same spatial grid as Fig. 3 for a piston with radius $a=5\lambda$. A single sector that covers the entire computational grid in Fig. 1 is defined for these computations. A comparison of the raw computation times indicates that, for the same number of Gauss abscissas, the times required to compute the pressure field using the expressions in H_2, H_3, and H_4 are roughly the same, although H_3 (dot dash) is somewhat slower because of the additional complex term in the numerator of Eq. (7). Likewise, for the source geometry and computational grid evaluated here, the expressions H_2, H_3, and H_4 are all nearly three times faster than the impulse response solution H_1 (dotted line).

pression, the entire near field is computed with one Gauss abscissa, then two Gauss abscissas, and so on up to 200. The resulting peak errors $\eta_{\max}$ are then plotted for calculations performed in a single sector ($n_s=1$) as a function of the number of Gauss abscissas applied to H_1, H_2, H_3, and H_4.

Figure 4 shows that, for a given number of Gauss abscissas, the efficient formulation H_4 (solid line) produces the smallest peak errors $\eta_{\max}$ overall. The peak errors are reduced relative to H_1, H_2, and H_3 because a singularity is subtracted from the numerator in H_4 and because the limits of integration are selected such that all redundancies are eliminated from the integral expression in H_4. If the near-field pressure distribution is computed with a constant number of Gauss abscissas at all grid points as in Fig. 4, the peak errors obtained with the Archer-Hall *et al.*[6] formulation H_2 (dashed line) are likewise relatively small. This result suggests that, for computations using a constant number of Gauss abscissas, the errors produced by the singularity in H_2 are dominated by errors in other locations. In this case, the largest error values are encountered in locations where $\theta=\sin^{-1}(r/\sqrt{r^2+z^2})$ is large. However, if the region near the radiating surface of the piston is excluded, the largest error values occur wherever $r\approx a$ as a result of the singularity. In Fig. 4, the peak errors obtained with the Hutchins *et al.*[7] formulation H_3 (dot dash) are consistently larger than those obtained with H_2 and H_4 because H_3 doubles the range over which the integrand is evaluated. The integral in H_3 is evaluated over the range $[-\pi,\pi]$, but the integrand is an even function, so the abscissas within the interval $[-\pi,0]$ are redundant. This redundancy increases the error relative to H_4 and H_2 because H_3 samples the integrand half as often within the range $[0,\pi]$. Figure 4 also demonstrates that the errors obtained with the impulse response H_1 (dotted line) are similar to those achieved with the efficient method H_4 over a very small range of values. However, the peak errors $\eta_{\max}$ are larger with the impulse response H_1 than any of the other single integral expressions for most values. The errors produced by the impulse response are the result of a rapid change in the impulse response where $r\approx a$.

Figure 5 illustrates the computation times, plotted as a function of the number of Gauss abscissas, that are associated with the peak errors depicted in Fig. 4. In Fig. 5, the same number of Gauss abscissas are applied throughout the entire pressure field for calculations using the expressions H_1, H_2, H_3, and H_4, and the computation times are measured for each pressure field distribution. Figure 5 shows that, if the number of Gauss abscissas remains constant throughout the computational grid depicted in Fig. 1, H_2 and H_4 compute near-field pressure distributions in the shortest times by almost a factor of 3. This result evaluates only the raw computation times for the same number of Gauss abscissas without considering the peak error $\eta_{\max}$. In Fig. 5, the Hutchins *et al.*[7] formulation H_3 is somewhat slower because of the additional complex term in the numerator of Eq. (7). Nevertheless, the expressions H_2, H_3, and H_4 are all nearly three times faster than the impulse response solution H_1 for the $a=5\lambda$ circular source and computational grid evaluated here. The impulse response H_1 is slower than the other expressions in part because of the additional time required to compute the inverse trigonometric function in Eq. (4) and in part because the other three expressions exploit the values that are repeated within the limits of integration and the integrand.

If the computation times in Fig. 5 are normalized with

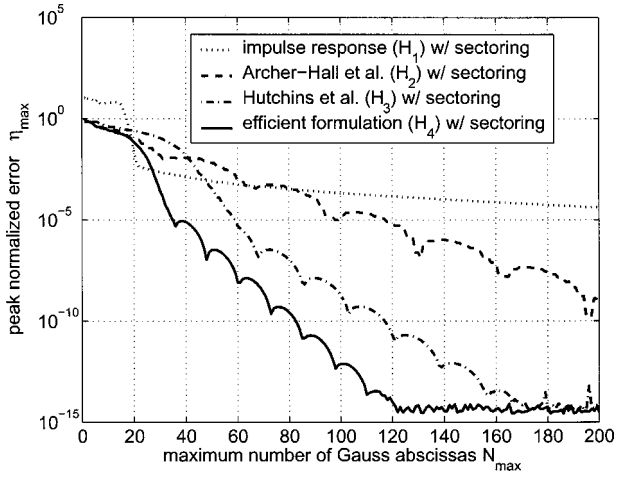


FIG. 6. A comparison of the maximum normalized errors $\eta_{\max}$ for beam patterns computed with H_1, H_2, H_3, and H_4 and the grid sectoring approach defined in Eqs. (11) and (12) with $n_s=10$ sectors. Each beam pattern is computed for a piston radius $a=5\lambda$ and compared to the reference beam pattern in Fig. 3. Peak error values $\eta_{\max}$ are plotted as a function of the maximum number of Gauss abscissas $N_{\max}$. The number of Gauss abscissas M in each sector varies according to Eqs. (11) and (12) so that M ranges from $0.3N_{\max}$ to $N_{\max}$ across 10 sectors. Relative to the errors computed in Fig. 4 for single sector ($n_s=1$), the peak errors $\eta_{\max}$ are maintained when grid sectoring with $n_s=10$ is combined with the efficient formulation H_4 (solid line). For the same combination of input values, grid sectoring combined with H_3 (dot dash) generally results in unchanged peak error values outside of the range $60\geqslant N_{\max}\geqslant 40$. For $n_s=10$ and $N_{\max}>20$, a noticeable increase in the peak error $\eta_{\max}$ is indicated for sectored results computed with the impulse response H_1 (dotted line), and the peak errors $\eta_{\max}$ computed with H_2 (dashed line) increase dramatically.

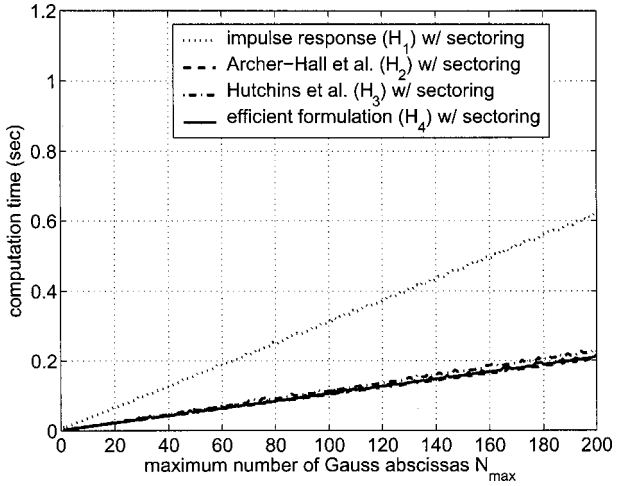


FIG. 7. A comparison of computation times for beam patterns calculated with H_1, H_2, H_3, and H_4 combined with the grid sectoring approach defined in Eqs. (11) and (12) and plotted as a function of the maximum number of Gauss abscissas $N_{\max}$. Each beam pattern is computed for a piston radius $a=5\lambda$ on the same spatial grid as the reference in Fig. 3. The number of Gauss abscissas M in an individual sector ranges from $0.3N_{\max}$ to $N_{\max}$ across the $n_s=10$ sectors indicated in Fig. 2. The axes are the same as in Fig. 5, so grid sectoring achieves roughly a factor of 2 reduction in the raw computation time for each integral formulation evaluated on this computational grid. Since grid sectoring reduces the computation time for each integral expression by nearly a factor of 2, H_2, H_3, and H_4 are again approximately three times faster than the impulse response solution H_1.

respect to the peak error values in Fig. 4, the resulting reduction in computation time achieved by the efficient formulation H_4 depends on the value of the specified peak value $\eta_{\max}$. This normalized value, evaluated in a single sector that encompasses the entire computational grid depicted in Fig. 1, compares the numerical performance of H_4 with that of the other single integral expressions H_1, H_2, and H_3 and summarizes the result with a single number. For example, the efficient formulation H_4 first achieves a peak error value below 10% in Fig. 4 with 19 Gauss abscissas, and the impulse response H_1 first reaches a peak error less than 10% in the same figure with 20 Gauss abscissas. The computation time for the efficient formulation H_4 using 19 Gauss abscissas with $n_s=1$ is then divided into the computation time for the impulse response using 20 Gauss abscissas with $n_s=1$, and the result indicates that the efficient method in H_4 is 3.2 times faster than the impulse response H_1 for a specified peak error value of 10%. For the same specified peak error, H_2 and H_4 are computed in roughly the same time (H_2 is marginally faster), whereas H_4 is 1.7 times faster than H_3. Similarly, for a peak error value of 1%, H_4 is 2.5 times faster than H_1, H_4 is roughly the same speed as H_2 (H_2 is again marginally faster), and H_4 is again 1.7 times faster than H_3. Finally, for a peak error value of 0.1%, H_4 is 3.6 times faster than H_1, H_4 is roughly the same speed as H_2, and H_4 is once again 1.7 times faster than H_3. These ratios are specific to the grid and source geometry considered here, and any changes in the size of the piston or the extent of the computational grid are expected to modify all of the calculated values.

C. Errors and times for multiple sectors ($n_s=10$)

The results of the grid sectoring strategy outlined in Eqs. (11) and (12), applied to a piston with $a=5\lambda$ and $n_s=10$, are summarized in Figs. 6 and 7. Figure 6 contains the peak computed errors for H_1, H_2, H_3, and H_4, and Fig. 7 displays the corresponding computation times. Figure 6 shows that, with 10 sectors defined, the efficient formulation H_4 (solid line) again achieves the smallest peak errors $\eta_{\max}$ overall. The peak errors computed for 10 sectors with H_4 in Fig. 6, compared to the peak errors computed with ($n_s=1$) in Fig. 4, are virtually unchanged for all values of $N_{\max}$. Thus, grid sectoring successfully maintains the peak error value for H_4 by applying the largest number of Gauss abscissas in the sector adjacent to the piston face and an incrementally smaller number of Gauss abscissas in successive sectors. For H_4 and the remaining single integral expressions, the peak errors $\eta_{\max}$ obtained with the maximum number of Gauss abscissas in Fig. 6 are compared to the peak errors $\eta_{\max}$ obtained with a constant number of Gauss abscissas in Fig. 4. With respect to the single sector ($n_s=1$) results presented in Fig. 4 for the same source and grid geometry, grid sectoring results with $n_s=10$ applied to H_3 (dot dash) generally produce the same peak error values $\eta_{\max}$ for all values of $N_{\max}$ except $60\geqslant N_{\max}\geqslant 40$. In contrast, the impulse response H_1 (dotted line) evaluated in $n_s=10$ sectors increases the peak errors $\eta_{\max}$ by a factor of 2 or more for all values of $N_{\max}>20$. Furthermore, the peak errors $\eta_{\max}$ computed for the Archer-Hall *et al.*[6] formulation H_2 increase by several orders of magnitude for values of $n_s=10$ and

$N_{max}>20$. With $n_s=10$, the peak error values in H_1 and H_2 are significantly increased, whereas H_3 and the efficient formulation H_4 maintain similar or equal peak error values, respectively.

Figure 7 shows that, for a circular piston with $a=5\lambda$ evaluated on the grid depicted in Fig. 1, grid sectoring with $n_s=10$ applied to H_1, H_2, H_3, and H_4 cuts the raw computation time in half. The times in Fig. 7 correspond to the errors in Fig. 6, where the values illustrated in these two figures are obtained from the same set of calculations. In Fig. 7, the grid and source geometries are the same as in previous figures, and the measured times in Fig. 7 are plotted on the same axes as Fig. 5 for purposes of comparison. The measured times for each expression in both Fig. 5 and Fig. 7 are linearly proportional to the number of Gauss abscissas and the maximum number of Gauss abscissas, respectively. As in Fig. 5, the time required to compute pressure fields with the expressions H_2, H_3, and H_4 is reduced by almost a factor of 3 in Fig. 7 relative to time required for impulse response calculations with H_1.

If the computation times in Fig. 7 are normalized with respect to the peak error values in Fig. 6, the reduction in computation time achieved by the efficient formulation H_4 again depends on the value of the specified peak value η_{max}. This normalized value, evaluated for $n_s=10$, compares the numerical performance of H_4 with that of the other single integral expressions H_1, H_2, and H_3 and summarizes the result for grid sectoring with a single number. For example, the efficient formulation H_4 initially achieves a peak error value below 10% in Fig. 6 where $N_{max}=19$, and the impulse response H_1 first reaches a peak error less than 10% in the same figure where $N_{max}=20$. The computation time for the efficient formulation H_4 with $n_s=10$ and $N_{max}=19$ is then divided into the computation time for the impulse response H_1 with $n_s=10$ and $N_{max}=20$, and the result indicates that the efficient method in H_4 is 3.1 times faster than the impulse response H_1 for a specified peak error value of 10%. For the same specified peak error, H_2 and H_4 are computed in roughly the same time (now with H_4 marginally faster), whereas H_4 is 1.7 times faster than H_3. For a peak error value of 1%, H_4 is 2.4 times faster than H_1, H_4 is 1.7 times faster than H_2 (H_4 and H_2 are no longer similar), and H_4 is again 1.7 times faster than H_3. Finally, for a peak error value of 0.1%, H_4 is 4.7 times faster than H_1, H_4 is 2.1 times faster than H_2, and H_4 is 1.7 times faster than H_3. Once again, these values are for a specific grid and source geometry, and any modifications in the grid or source are expected to change these computed values.

These normalized computation times show that, if grid sectoring is applied, the efficient formulation H_4 computes the near-field distribution more quickly than any of the other single integral expressions. This result is expected, since the peak errors and the computation times are smallest with the efficient formulation, both for results computed in a single sector (Figs. 4 and 5) and in 10 sectors (Figs. 6 and 7). However, the relationship between the peak errors computed with H_4 and H_2 changes dramatically with 10 sectors, and for peak errors of 1% and smaller, the normalized computation time required for the Archer-Hall *et al.*[6] expression H_2 acutally increases with $n_s=10$. Thus, grid sectoring is not recommended for pressure field calculations with H_2 that require errors less than 1%, although grid sectoring reduces the normalized computation time for all of the other expressions (H_1, H_3, and H_4).

D. Spatial distribution of the error

If a single sector ($n_s=1$) is defined for near-field pressure calculations, the spatial distribution of the error $\eta(r,z)$ changes for each single integral expression as the number of Gauss abscissas is increased. Initially, the largest errors are increasingly concentrated away from the piston normal and toward the piston face (where $\theta=\sin^{-1}\{r/\sqrt{r^2+z^2}\}$ is largest) until the complex exponential terms in each integrand are adequately sampled. For a circular piston with $a=5\lambda$ evaluated across the range of coordinate values depicted in Fig. 1, the large errors adjacent to the piston face disappear if near-field pressures are evaluated with 20 or more Gauss abscissas are applied to H_1, H_2, and H_4 and if more than 30 Gauss abscissas are applied to H_3. When H_1 and H_2 are evaluated with additional Gauss abscissas, numerical singularities appear in locations where $r\approx a$. Examples of the spatial distribution of the error $\eta(r,z)$ for a single sector (not shown) are easily obtained for each of the single integral expressions in H_1, H_2, H_3, and H_4.

Illustrations of $\eta(r,z)$ evaluated in multiple sectors ($n_s=10$) are demonstrated in Figs. 8–11. These figures show that the expressions H_1, H_2, H_3, and H_4 produce four distinct error distributions and that grid sectoring with $N_{max}=50$ has a different effect on each expression. In Figs. 8–11, the error distribution in the sector adjacent to the piston face is the same as that obtained if $N_{max}=50$ is applied throughout the computational grid. The error distributions in the remaining sectors differ from those obtained with $N_{max}=50$ because the number of Gauss abscissas decreases in successive sectors. The error distributions $\eta(r,z)$ in Figs. 8–10 demonstrate that the peak values η_{max} for H_1, H_2, and H_3 are located in sectors away from the piston face, which indicates that these three expressions increase the peak error value relative to that obtained when N_{max} Gauss abscissas are applied throughout the entire grid. If the peak error value η_{max} occurs in the sector adjacent to the piston face (where N_{max} Gauss abscissas are applied) as in Fig. 11, then the peak error value is the same as that obtained when N_{max} Gauss abscissas are applied throughout the entire grid. In this example, H_4 maintains the peak error with values of $n_s=10$ and $N_{max}=50$, while the expressions H_1, H_2, and H_3 increase the peak error with the same parameters.

In Fig. 8, the normalized error $\eta(r,z)$ is computed for the impulse response H_1 evaluated with $n_s=10$ sectors and $N_{max}=50$ Gauss abscissas for a circular piston with radius $a=5\lambda$. Figure 8 contains an example of the numerical singularity produced by the impulse response which appears adjacent to the line $r/a=1$. This singularity is observed well beyond the grid boundaries depicted in these figures. The numerical artifact in this region is caused by the slope of the $\cos^{-1}\{(\beta^2-z^2+r^2-a^2)/(2r\sqrt{\beta^2-z^2})\}$ term in the impulse response H_1, which is infinite at the values for β_2 and β_3

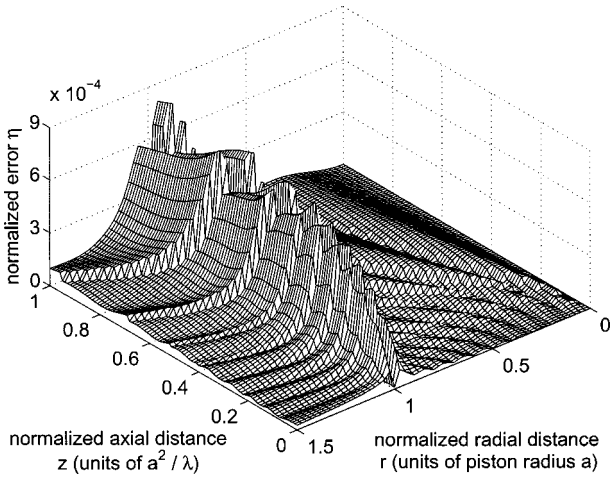


FIG. 8. Normalized difference between the reference beam pattern in Fig. 3 and the results of the impulse response H_1 evaluated with grid sectoring as defined by Eqs. (11) and (12). For this simulation, the near-field region is divided into $n_s=10$ sectors, and the scale factors $C=5$ and $S=0.3$ specify the largest and smallest number of abscissas as 50 and 15, respectively. The integrand is evaluated with the largest number of Gauss abscissas in the sector adjacent to the piston face and the smallest number of Gauss abscissas in the sector adjacent to the piston normal. When these values are applied to H_1 for a circular piston with radius $a=5\lambda$, the maximum normalized error is $\eta_{\max}=8.33\times10^{-4}$ relative to the reference in Fig. 3. As the number of Gauss abscissas decreases in each sector evaluated along the line $r/a=1$, the computed numerical error grows larger with increasing values of z. Numerical errors are encountered along the line $r/a=1$ because of the large slopes produced by the $\cos^{-1}$ term in the impulse response.

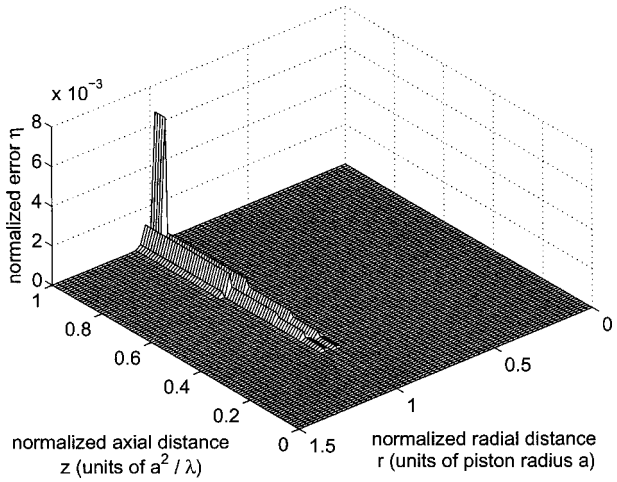


FIG. 9. Normalized difference between the reference beam pattern in Fig. 3 and the results of the Archer-Hall *et al.* (Ref. 6) formula H_2 evaluated with grid sectoring as defined by Eqs. (11) and (12). For this simulation, the near-field region is divided into $n_s=10$ sectors, and the scale factors $C=5$ and $S=0.3$ specify the largest and smallest number of abscissas as 50 and 15, respectively. The integrand is evaluated with the largest number of Gauss abscissas in the sector adjacent to the piston face and the smallest number of Gauss abscissas in the sector adjacent to the piston normal. When these values are applied to H_2 for a circular piston with radius $a=5\lambda$, the maximum normalized error is $\eta_{\max}=6.63\times10^{-3}$ relative to the reference in Fig. 3. The errors shown here grow substantially larger with increasing z near the line $r/a=1$. These errors are produced by a numerical singularity in H_2.

specified in Eq. (2). These infinite slopes are encountered in calculations of the impulse response at all grid coordinates, and the errors near line $r/a=1$ are amplified by other rapid changes that are caused by the $\cos^{-1}$ term in I_1. In Fig. 8, the errors caused by the singularity grow larger as the value of z increases, and several discrete jumps are evident along the line $r/a=1$. These jumps, which appear at the sector boundaries defined by Eq. (11), occur as the number of Gauss abscissas applied within each sector decreases. In subsequent sectors encountered as the z coordinate increases along the line $r/a=1$, the errors rise rapidly at each sector boundary, then decay until the next sector boundary is reached. This pattern is repeated in subsequent sectors. The error distribution in Fig. 8 suggests that additional Gauss abscissas are required near $r/a=1$, whereas grid sectoring reduces the number of Gauss abscissas with increasing z values in this location. Grid sectoring nevertheless reduces the normalized computation time for the impulse response H_1 evaluated on the computational grid in Fig. 1, though the improvements diminish for values of $N_{\max}>20$.

Figure 9 depicts the error distribution $\eta(r,z)$ computed for the Archer-Hall *et al.*[6] expression H_2 with $n_s=10$ sectors and $N_{\max}=50$ Gauss abscissas. Figure 9 contains an example of $\eta(r,z)$ calculated with H_2 for a circular piston with radius $a=5\lambda$ where the error increases along the line $r/a=1$ as z coordinate moves away from the face of the piston. This error distribution, which produces larger errors as the number of samples is reduced by grid sectoring, is the result of the numerical singularity in H_2. The singularity originates in I_2, where the denominator of the integrand in Eq. (6) approaches zero as the value of r approaches a. As with the impulse response H_1, the numerical error computed with the expression H_2 jumps at each sector boundary encountered along $r/a=1$ as z increases, suggesting that H_2 requires additional abscissas in this location. However, grid sectoring reduces the number of Gauss abscissas in each subsequent sector, and the errors computed with H_2 increase where $r\approx a$. As a result, the singularity in H_2 limits the effectiveness of grid sectoring in Fig. 1, especially for values of $N_{\max}>20$.

Figure 10 shows the spatial distribution of $\eta(r,z)$ for the Hutchins *et al.*[7] expression H_3 with $n_s=10$ sectors and $N_{\max}=50$ Gauss abscissas. Unlike the error distributions presented in Figs. 8 and 9 for H_1 and H_2, respectively, the errors depicted in Fig. 10 for a circular source with $a=5\lambda$ demonstrate that H_3 eliminates the singularity along the line $r/a=1$. However, as the value of z increases in successive sectors and as the number of Gauss abscissas is therefore reduced, the peak value of the error observed in each sector gradually rises at first until a maximum value is reached and then slowly diminishes. This error distribution indicates that the linear mapping defined for $n_s=10$ sectors in Eq. (12) underestimates the number of Gauss abscissas required for calculations with H_3 within intermediate sectors, particularly for values of $60\geqslant N_{\max}\geqslant 40$. Outside of this range of values for $N_{\max}$, where the exact range of values for $N_{\max}$ is a function of the source and grid parameters, the peak errors consistently decrease from one sector to the next as z increases. Although Fig. 10 indicates that, for certain parameters, H_3 increases the numerical error obtained with grid sectoring, the overall numerical performance of H_3 is improved by subtracting the singularity and eliminating the numerical problems along the line $r/a=1$.

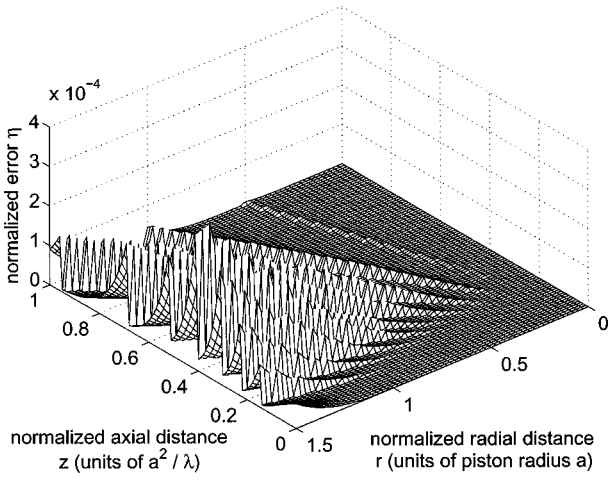


FIG. 10. Normalized difference between the reference beam pattern in Fig. 3 and the results of the Hutchins *et al.* (Ref. 7) formula H_3 evaluated with grid sectoring as defined by Eqs. (11) and (12). In this simulation, the near-field region is divided into $n_s=10$ sectors, and the scale factors $C=5$ and $S=0.3$ specify the largest and smallest number of abscissas as 50 and 15, respectively. The integrand is evaluated with the largest number of Gauss abscissas in the sector adjacent to the piston face and the smallest number of Gauss abscissas in the sector adjacent to the piston normal. When these values are applied to H_3 for a circular piston with radius $a=5\lambda$, the maximum normalized error is $\eta_{\max}=3.67\times10^{-4}$ relative to the reference in Fig. 3. Similar to the result in Fig. 8, the sectoring strategy increases the maximum error value somewhat relative to the peak value obtained if the results are evaluated with 50 Gauss abscissas throughout the field; however, no singularity is present along the line $r/a=1$ because a singularity is subtracted within the integrand of H_3.

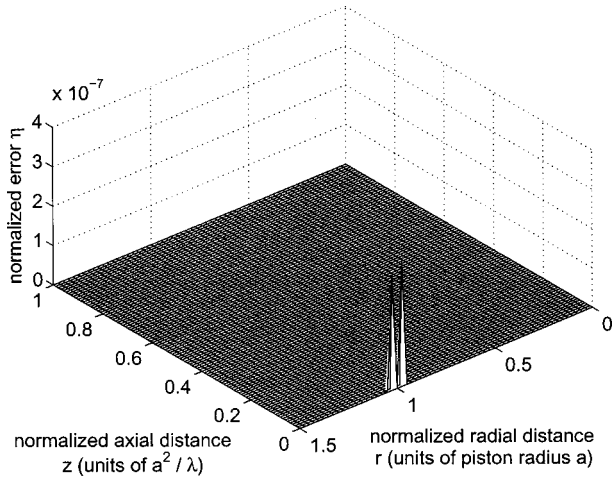


FIG. 11. Normalized difference between the reference beam pattern in Fig. 3 and the results of the numerically efficient formulation H_4 evaluated with grid sectoring as defined by Eqs. (11) and (12). In this simulation, the near-field region is divided into $n_s=10$ sectors, and the scale factors $C=5$ and $S=0.3$ specify the largest and smallest number of abscissas as 50 and 15, respectively. The integrand is evaluated with the largest number of Gauss abscissas in the sector next to the piston face and the smallest number of Gauss abscissas in the sector adjacent to the piston normal. When these values are applied to H_4 for a circular piston with radius $a=5\lambda$, the maximum normalized error is $\eta_{\max}=3.20\times10^{-7}$ relative to the reference in Fig. 3. As in Fig. 10, and in contrast to the results shown in Figs. 8–10, the sectoring strategy maintains a maximum error that is the same as that obtained if 50 Gauss abscissas are applied to all points in the field. The maximum error value is maintained because the limits of integration are minimized and a singularity is subtracted from H_4.

Figure 11 contains the distribution of computed error values $\eta(r,z)$ for the efficient formulation H_4 evaluated with $n_s=10$ and $N_{\max}=50$. Figure 11 shows that, for this combination of values evaluated on the grid in Fig. 1, the errors obtained with H_4 are substantially smaller than those produced by H_1, H_2, and H_3 in Figs. 8, 9, and 10, respectively. In Fig. 11, the computed errors are confined to a small region where $z\approx0$ and $r\approx a$. Although some errors remain after the numerical singularity of H_2 shown in Fig. 9 is removed in H_4, the remaining errors in Fig. 11 are insignificant. Likewise, by selecting limits of integration that remove all redundant function evaluations, the efficient formulation also avoids the numerical problems illustrated in Fig. 10. Figure 11 shows that, although $n_s=10$ sectors are defined for H_4, the sector boundaries observed in previous mesh plots are not present. The sector boundaries disappear from the plot of $\eta(r,z)$ whenever the errors in subsequent sectors are significantly smaller than the errors in the sector adjacent to the circular source. Thus, for grid sectoring defined with $n_s=10$ and $N_{\max}=50$, the efficient formulation H_4 maintains the same peak error value obtained when 50 Gauss abscissas are applied throughout the computational grid. A comparison of Figs. 4 and 6 indicates that the efficient expression H_4 maintains the same peak error for all values of $N_{\max}$.

V. DISCUSSION

A. Single sector

The single integral expressions H_1, H_2, H_3, and H_4 are all analytically equivalent; however, the numerical properties of each vary widely. For computations that define a single sector ($n_s=1$) so that the number of Gauss abscissas remains constant across the entire computational grid, the Archer-Hall *et al.*[6] expression H_2 and the efficient expression H_4 produce pressure fields with approximately the same peak errors in roughly the same amount of time. This result is restricted to near-field pressures calculated in the grid depicted in Fig. 1. For other grid geometries that include points with $r\approx a$ and exclude the region near the surface of the circular piston, the peak error obtained with H_2 increases relative to that obtained with H_4. Therefore, the normalized computation time for H_2 evaluated in a single sector is approximately the same as that for the efficient formulation H_4 only in certain circumstances. The normalized computation times for H_4 are consistently less than or equal to those for H_2, so the overall numerical performance of H_4 is superior to that of H_2, even for calculations performed within a single sector.

As indicated in Figs. 4 and 5, the efficient formulation H_4 consistently produces smaller errors in less time than the Hutchins *et al.*[7] expression H_3. The normalized computation times, which condense these results into a single value, show that H_4 is 1.7 times faster than H_3 for specified errors of 10%, 1%, and 0.1%, where the exact value of this ratio varies as a function of the grid and source geometry. For calculations in a single sector, H_3 consistently requires more time than H_4 to achieve a specified peak error. This is due to the additional term in the numerator of Eq. (7), which increases the raw computation time, and the duplicated values in the

integrand, which increase the error relative to that obtained with H_4.

Figure 5 also shows that, for pressure field calculations performed in a single sector that covers the entire grid in Fig. 1, the computation times measured with the impulse response H_1 are significantly larger than those time than those obtained with the same number of Gauss abscissas applied to the efficient formulation H_4. Although the peak errors produced by H_1 are slightly smaller than those generated by H_4 if the number of Gauss abscissas falls within a certain restricted range, comparisons of normalized computation times demonstrate that the efficient formulation H_4 is consistently faster than the impulse response H_1. In calculations performed on a 61 by 101 point grid, H_4 is 3.1, 2.4, and 4.7 times faster than H_1 for peak error values of 10%, 1%, and 0.1%, respectively. In computational grids that contain a larger number of points, the efficient formulation H_4 is even faster relative to H_1 because calculations involving common terms in Eq. (8) are exploited in even more locations.

The impulse response H_1 consistently requires more time for pressure field calculations than the other three single integral formulations because the values for the limits of integration β_2 and β_3 are constantly changing and because the only expression repeated in I_1 is (β^2-z^2). By computing new values for β_2 and β_3 at every new value of r and z, the computational overhead is increased relative to that required for H_2, H_3, and H_4. In contrast, the other three single integral expressions compute the limits of integration once and then repeatedly apply the result at every point in the computational grid. H_2, H_3, and H_4 also share some other repeated expressions that are very similar. These three single integral formulations, which all include integrals that are evaluated with respect to ψ, contain the term $r^2+a^2-2ar\cos\psi$ in both the denominator and an exponent of the integrand. This term is independent of the z coordinate and is therefore computed once for each value of r and then applied repeatedly at each z value that shares the same radial coordinate. Similarly, H_2, H_3, and H_4 each contain ratios of terms that are strictly functions of r, a, and ψ and therefore independent of z. These expressions are also evaluated once for each value of r and then applied repeatedly at each z coordinate. In addition, the value of e^{-jkz} is calculated once for each z coordinate in the grid and then applied throughout the computational grid wherever that z value is encountered. Even without these time-saving measures, pressure field calculations are completed more quickly with H_2, H_3, and H_4 than with the impulse response H_1 because of the extra time required to compute the inverse cosine function. These improvements apply equally to computations with a single sector and computations with multiple sectors, as demonstrated by the consistent relationships between computation times demonstrated in Figs. 5 and 7.

B. Multiple sectors

In pressure fields computed with multiple sectors, the Archer-Hall *et al.*[6] expression H_2 produces larger errors than the efficient formulation H_4. This is clearly demonstrated in a comparison of the peak errors $\eta_{\max}$ depicted in Fig. 6. For calculations with $N_{\max}>20$ and $n_s=10$ in a 61 by 101 point grid, the peak errors computed with H_2 are several orders of magnitude greater than those computed with H_4. Although the specific error values change with the number of sectors and the value of $N_{\max}$, the peak errors produced by H_2 are consistently larger than those obtained with H_4 whenever the pressure field is sampled near $r/a=1$. In this region, the errors grow larger as the sampling of the integrand is reduced because of the singularity in H_2, and H_4 avoids these errors by subtracting the singularity in Eq. (8). The results indicate that H_2 and H_4 achieve an error of 10% in approximately the same amount of time; however, this result only applies to the grid geometry in Fig. 1. Other grid geometries that evaluate the pressure at some distance from the surface of the circular piston show that H_4 is faster than H_2 unless values near $r/a=1$ are avoided completely. The ratios of normalized computation times corresponding to peak errors of 1% and 0.1% measured for the same grid show that H_4 is 1.7 and 2.1 times faster than H_2, respectively. These values also depend on the grid and source geometry. Relative to results computed with a single sector ($n_s=1$) defined for the entire grid in Fig. 1, the normalized computation time for H_2 actually increases with $n_s=10$ sectors for specified peak errors less than or equal to 0.1%. This result indicates that grid sectoring applied to H_2, by reducing the number of Gauss abscissas in sectors away from the face of the circular piston, shifts samples away from locations where a singularity dominates the peak error values. Grid sectoring therefore is primarily applicable either to grid coordinates that avoid the singularities or integral formulations that eliminate singularities in all coordinate locations.

A comparison between the expression H_3 derived by Hutchins *et al.*[7] and the efficient formulation H_4 shows that expressions that subtract the singularity obtain consistent numerical results. Whether the specified peak error is 10%, 1%, or 0.1%, the efficient expression H_4 evaluated on a 61 by 101 point grid defined in Fig. 1 is 1.7 times faster than H_3. The ratio of normalized computation times changes as the source and grid parameters vary, but the results are much more consistent for different error values computed with the same source and grid parameters if the singularity is subtracted as in H_3 and H_4.

Relative to the peak errors $\eta_{\max}$ obtained from near-field pressure calculations that apply the same number of Gauss abscissas to the impulse response H_1 at every grid point in Fig. 1, the peak errors increase for H_1 with $n_s=10$ sectors whenever $N_{\max}>20$. As shown in a comparison of the peak errors for H_1 computed with one and ten sectors in Figs. 4 and 6, respectively, the peak error obtained with H_1 increases by at least a factor of 2 up to a factor of 8 for this range of $N_{\max}$ values. As with all of the other results, the exact ratios are a function of the source and grid parameters. The peak error values for H_1 in Fig. 6, which are calculated with $n_s=10$ sectors, are larger that those computed with the Archer-Hall *et al.*[6] expression H_2 for values of $N_{\max}>80$, where both peak errors are about 0.03%. In terms of normalized computation times, H_2 is 3.0, 1.4, and 2.2 times faster than H_1 for specified peak errors of 10%, 1%, and 0.1%, respectively, for pressure field calculations evaluated in $n_s=10$ sectors. Therefore, even with $n_s=10$ sectors defined for

grid sectoring, the impulse response is the slowest of the four single integral expressions for this combination of parameters. Overall, for near-field calculations performed in single or multiple sectors, the fastest and most consistent performance is demonstrated by the efficient method H_4 in Eq. (8).

C. Parameters for grid sectoring

Each of the near-field pressure distributions computed with grid sectoring defines $n_s=10$ sectors for the grid outlined in Fig. 1. For this grid, any value of n_s between 5 and 15 achieves approximately the same factor of 2 reduction in the computation time demonstrated in a comparison between Figs. 5 and 7. With fewer sectors, errors at the sector boundaries can exceed the peak error in the sector next to the surface of the radiating source, and values of $n_s>15$ reduce the number of repeated calculations in each sector for H_2, H_3, and H_4. Outside of this range of values for n_s, grid sectoring achieves some reduction in the normalized computation time, but the values $15\geqslant n_s\geqslant 5$ demonstrate the best performance for the computational grid evaluated here. The specific range of values depends on the source and grid geometry, so grid sectoring is defined for different values of n_s with each new combination of parameters.

The linear mapping function defined in Eq. (12) eliminates large errors that otherwise appear in sectors located near the piston normal. Although near-field pressure calculations in the axial region require fewer Gauss abscissas than other nearfield locations, Eq. (9) underestimates the number of Gauss abscissas required near the piston normal. For example, with $N_{\max}=10$ and $n_s=10$, Eq. (9) correctly employs 10 Gauss abscissas in the sector adjacent to the piston face, but Eq. (9) also applies one Gauss abscissa in the sector adjacent to the piston normal. If one Gauss abscissa is applied throughout this sector, the resulting errors are much larger than those obtained in the sector adjacent to the piston face. To prevent the errors in the axial region from increasing the peak error, the number of Gauss abscissas is computed with Eq. (12) instead of Eq. (9). In this example, Eq. (12) with $S=0.3$ employs three Gauss abscissas throughout the sector adjacent to the piston face, and additional abscissas are specified in subsequent sectors until the maximum value $N_{\max}=10$ is reached. The value $S=0.3$ is used for grid sectoring with each of the single integral expressions evaluated in Figs. 6–11; however, other values are required for different grid and source geometries.

D. Near-field/far-field transition

The pressure field in Fig. 3 and the error distributions in Figs. 8–11 are evaluated strictly in the near-field region; however, the single integral expressions H_1, H_2, H_3, and H_4 are also useful for pressure calculations well beyond the near-field/far-field transition distance. As demonstrated in the text by Kinsler *et al.*,[13] the far-field approximation for a circular piston produces substantial errors at the far-field transition distance a^2/λ. These errors continue well beyond the location of the near-field/far-field transition, out to at least 5 transition distances. The errors are characterized by discrete jumps in the computed pressure that are clearly evident on a linear scale that is normalized to the peak value of the axial pressure. On this scale, these jumps are no longer visible after six or seven transition distances for most circular piston geometries. Thus, the far-field approximation is only appropriate for locations that are at least six or seven times the transition distance a^2/λ. The efficient formulation H_4 combined with grid sectoring is also effective well beyond six or seven transition distances, which eliminates the need for two different calculation methods in grids that extend into both regions.

E. Large aperture phased arrays

The single integral calculations H_1, H_2, H_3, and H_4 are all substantially faster than any method that superposes point sources (e.g., Zemanek[8]) or rectangular radiators (e.g., Ocheltree and Frizzell[14]). In particular, near-field calculations that combine the efficient formulation H_4 with grid sectoring can reduce the computation time by as much as an order of magnitude or more relative to point source or rectangular radiator calculations. Shorter run times are especially useful in calculations of three-dimensional fields generated by large ultrasound phased arrays designed for thermal therapy. For example, pressure fields are repeatedly calculated in evaluations of sparse ultrasound phased arrays[15–17] designed for noninvasive surgery. Pressure field computations for large sparse arrays are very time-consuming, and the efficient formulation H_4 combined with grid sectoring achieves a significant reduction in the computation time relative to these other methods.

An expression similar to H_4 is also available for square or rectangular sources,[18] and this expression is likewise faster than any calculation that uses point sources, rectangular radiators, or the impulse response. With or without grid sectoring, the efficient formulation H_4 derived for a circular piston is faster than any of these expressions for rectangular sources. Calculations utilizing H_4 in repeated pressure field calculations save time by evaluating a single one-dimensional integral instead of a two-dimensional integral or multiple one-dimensional integrals, and because H_4 reduces the computation time, this expression is recommended for preliminary designs of sparsely and densely populated therapy arrays.

VI. CONCLUSION

In a comparison of four analytically equivalent single integral expressions that describe the near-field pressure produced by a circular piston, the efficient formulation H_4 in Eq. (8) achieves the smallest numerical errors in the shortest time. The efficient formulation H_4, which eliminates redundant calculations and subtracts a singularity, exploits the best features of the Archer-Hall *et al.*[6] expression H_2 and the Hutchins *et al.*[7] expression H_3. The single integral expression H_4 is also amenable to grid sectoring, where the computational grid is divided into sectors, and the number of Gauss abscissas is adjusted from one sector to the next. In the sector adjacent to the piston face, the pressure is evaluated with the largest number of Gauss abscissas, and in the sector adjacent to the piston normal, pressure calculations

employ the smallest number of Gauss abscissas. Sector definitions are improved with a scheme that defines straight lines emanating from the center of the circular piston. Grid sectoring is also improved with a linear mapping function that avoids sampling problems in the axial region. Grid sectoring applied to the efficient formulation H_4 successfully reduces the computation time by nearly a factor of 2 without increasing the peak error value.

Near-field pressure distributions are evaluated on a 61 by 101 point grid for a circular piston with radius $a=5\lambda$, and the results obtained for the single integral expressions H_1, H_2, H_3, and H_4 are evaluated both in a single sector and in $n_s=10$ sectors. If all pressure field values are calculated in a single sector that extends up to the piston face or if the computational grid avoids all locations where $r\approx a$, the Archer-Hall *et al.*[6] expression H_2 and the efficient formulation H_4 obtain nearly the same errors in approximately the same amount of time. Otherwise, the peak error values calculated with H_2 are greater than those obtained with H_4. For calculations utilizing multiple sectors, the efficient formulation H_4 also demonstrates superior performance. For example, the peak error increases dramatically with $n_s=10$ sectors for the Archer-Hall *et al.*[6] expression H_2, and the singularity in H_2 prevents grid sectoring from achieving significant reductions in computation times that are normalized with respect to the peak errors. The peak errors obtained with the impulse response H_1 are also amplified in $n_s=10$ sectors, though grid sectoring nevertheless reduces the normalized computation time for H_1 because the peak errors increase by less than an order of magnitude relative to calculations performed in a single sector. The peak errors obtained with the Hutchins *et al.*[7] expression H_3 increase somewhat over a small range of values for $N_{\max}$ in calculations with $n_s=10$ sectors, and despite this moderate increase in the peak error, grid sectoring achieves a consistent reduction in the normalized computation time. Likewise, the peak errors evaluated with the efficient formulation H_4 are effectively unchanged in calculations with $n_s=10$ sectors, so grid sectoring combined with H_4 consistently reduces the normalized computation time. Results evaluated in a 61 by 101 point grid divided into $n_s=10$ sectors for a piston with radius $a=5\lambda$ show that pressures are calculated by the efficient formulation H_4 in about the same time as H_2 for a peak error value of 10%, whereas H_4 is 3.1 times faster than the impulse response H_1 and 1.7 times faster than H_3. For a peak error value of 1%, the efficient formulation H_4 is 2.4, 1.7, and 1.7 times faster than H_1, H_2, and H_3, respectively. For a peak error value of 0.1%, the efficient formulation H_4 is 4.7, 2.1, and 1.7 times faster than H_1, H_2, and H_3, respectively. As demonstrated by these ratios of normalized computation times, the impulse response H_1 is the slowest of these four single integral expressions, and the efficient formulation H_4 is the fastest. The ratios of normalized computation times change as the grid and source parameters vary; however, H_4 consistently achieves the best overall performance.

ACKNOWLEDGMENTS

This work was funded in part by a Biomedical Engineering Research Grant from the Whitaker Foundation, NIH Grant 5PO1 CA42745, and NIH Grant 1R01 CA093669. The authors would like to thank Jeremy D. Hoff of the Department of Radiation Oncology in the Duke University Medical Center for his helpful comments and suggestions.

[1] F. Oberhettinger, "On transient solutions of the 'baffled piston' problem," J. Res. Natl. Bur. Stand., Sect. B **65B**, 1–6 (1961).
[2] P. R. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," J. Acoust. Soc. Am. **49**, 1629–1638 (1971).
[3] J. C. Lockwood and J. G. Willette, "High-speed method for computing the exact solution for the pressure variations in the near field of a baffled piston," J. Acoust. Soc. Am. **53**, 735–741 (1973).
[4] G. R. Harris, "Transient field of a baffled planar piston having an arbitrary vibration amplitude distribution," J. Acoust. Soc. Am. **70**, 186–204 (1981).
[5] W. A. Verhoef, M. J. T. M. Cloostermans, and J. M. Thijssen, "The impulse-response of a focused source with an arbitrary axisymmetric surface velocity distribution," J. Acoust. Soc. Am. **75**, 1716–1721 (1984).
[6] J. A. Archer-Hall, A. I. Bashter, and A. J. Hazelwood, "Means for computing the Kirchhoff surface integral for a disk radiator as a single integral with fixed limits," J. Acoust. Soc. Am. **65**, 1568–1570 (1979).
[7] D. A. Hutchins, H. D. Mair, P. A. Puhach, and A. J. Osei, "Continuous-wave pressure fields of ultrasonic transducers," J. Acoust. Soc. Am. **80**, 1–12 (1986).
[8] J. Zemanek, "Beam behavior within the near field of a vibrating piston," J. Acoust. Soc. Am. **49**, 181–191 (1971).
[9] P. J. Davis and P. Rabinowitz, *Numerical Integration* (Academic, New York, 1975), pp. 139, 140.
[10] D. P. Orofino and P. C. Pedersen, "Multirate digital signal-processing algorithm to calculate complex acoustic pressure fields," J. Acoust. Soc. Am. **92**, 563–582 (1992).
[11] P. J. Davis and P. Rabinowitz, *Numerical Integration* (Academic, New York, 1975), pp. 73–77, 87–89, 369.
[12] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1972), pp. 887–889 and 916–919.
[13] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals of Acoustics*, 4th ed. (Wiley, New York, 2000), p. 179.
[14] K. B. Ocheltree and L. A. Frizzell, "Sound field calculation for rectangular sources," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **36**, 242–248 (1989).
[15] S. A. Goss, L. A. Frizzell, J. T. Kouzmanoff, J. M. Barich, and J. M. Yang, "Sparse random ultrasound phased array for focal surgery," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **43**, 1111–1121 (1996).
[16] L. R. Gavrilov and J. W. Hand, "A theoretical assessment of the relative performance of spherical phased arrays for ultrasound surgery," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 125–139 (2000).
[17] M. Pernot, J.-F. Aubry, M. Tanter, J.-L. Thomas, and M. Fink, "High power transcranial beam steering for ultrasonic brain therapy," Phys. Med. Biol. **48**, 2577–2589 (2003).
[18] R. J. McGough, "Rapid calculations of time-harmonic nearfield pressures produced by rectangular pistons," J. Acoust. Soc. Am. (to be published).

Real-time focusing using an ultrasonic one channel time-reversal mirror coupled to a solid cavity

N. Quieffin,[a] S. Catheline, R. K. Ing, and M. Fink
Laboratoire Ondes et Acoustique, ESPCI, University of ParisVII, U.M.R. C.N.R.S. 7587, 10 rue Vauquelin, 75231 Paris Cédex 05, France

(Received 29 August 2003; revised 4 December 2003; accepted 19 February 2004)

Focusing and beam steering is achieved by using a time-reversal process and a single transducer coupled to a solid cavity that is immersed in water. This low-cost technique makes it possible to focus acoustic energy anywhere on a 3D domain with a spatio-temporal resolution comparable to that of multiple transducers array. A short pulse is emitted from a transducer stuck at the surface of the solid cavity. The multiple-scattered field is measured in front of the solid cavity using a hydrophone needle at a reference point. This signal is then time reversed and remitted by the transducer. Around the reference point, one can observe a spatio-temporal recompression. The sidelobe level as well as the focal width no longer depend on the transducer aperture but on the dimensions of the solid cavity and the multiple paths covered by the acoustic waves in the solid. Moreover, it is shown how the experimental impulse responses on the front face of the cavity can be used to control the emitting ultrasonic field. This ''synthetic time-reversal'' technique is shown to be as powerful as a real time-reversal process.
[DOI: 10.1121/1.1699396]

PACS numbers: 43.20.Tb, 43.40.Hb, 43.60.Tj [EJS] Pages: 1955–1960

I. INTRODUCTION

To focus acoustic waves anywhere in a 3D domain, one can use a focused transducer with a mechanical scanning or a transducer array with an electronic beamforming. The first technique requires a very simple setup but isn't a real-time technique. The second one is a real-time technique but the electronic system becomes cumbersome for a 3D real-time device. Here, a low-cost and real-time method is proposed, that uses a single transducer coupled to a solid cavity and the time-reversal process.[1]

In previous works, Draeger[2–4] and J. De Rosny[5] have shown the feasibility of focusing ultrasounds anywhere inside a quasiclosed 2D cavity with only one source and the time-reversal process. With the time-reversal too, Derode[6] focuses ultrasounds with one transducer through open multiple-scattering medium. Our aim is to focus acoustic field in a 3D domain, outside the solid structure, just like a 2D array, with only one transducer. As it is shown in Fig. 1, a radiating 3D solid cavity is used. The use of a single source (instead of an array) involves a lack of spatial information. When this source is coupled to the cavity, the lost information is compensated by an increase of temporal information. The time-reversal focusing experiment comprises two stages. The first one is the acquisition of the impulse response: a source, located in the water, emits a short pulse which propagates toward the cavity and reflects inside. The echoes are collected by a contact transducer working as a receiver. The duration of the response just depends on the absorption of the material and on the energy radiation property of the cavity. During the second stage, the contact transducer emits the time-reversed version of the latter signal so that a part of the emitted acoustic field reflects inside the cavity, following the inverse way of propagation, and focuses at the point of first emission in the water. To focus in another position, the source in the water is moved and the same process is applied. Finally, a ''library'' of reference signals is constructed in order to be able to focus anywhere in a 3D domain.

First, we present a simple mathematical description of the time-reversal process involving a cavity. It is shown that the spatial dimensions of the focusing spot no longer depend on the aperture of the transducer but on the cavity's dimensions. Similar results are observed in a multiple-scattering medium: Derode[6] uses a 1 or 128 transducers array and time-reversal process to focus ultrasound through a random distribution of steel rods. The spatial resolution of the system is enhanced by the multiple scattering compared to the resolution of the array in the water. Indeed, the spatial high frequencies, which are ordinarily lost during the propagation, are redirected to the array because of the multiple scattering, so that the angular aperture of the system is larger than the array alone. These results are also observed in the underwater acoustic by Root and Rogers,[7] by using a 27-element volume array where different scattering structures are present within the volume of the array. In our case, the effective aperture of the system is defined by the surface of the cavity in front of the domain where we want to focus. Moreover, this definition of the aperture agrees with the diffraction laws.

A deeper analysis of the mathematical diffraction theory leads us to greatly simplify the construction of the library of reference signals. Indeed, the impulse responses need not be acquired on a 3D domain. According to the Helmholtz integral theorem they can be deduced from the impulse response acquired on a 2D aperture: the front face of the cavity. Therefore, one can compute a time-reversal signal able to focus on

[a]Electronic mail: nicolas.quieffin@loa.espci.fr

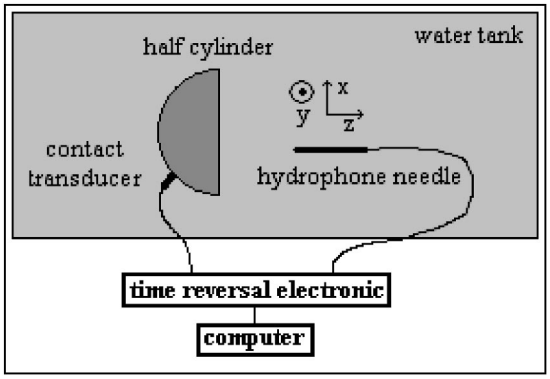


FIG. 1. Top view of the experimental setup. The contact transducer is fixed to the duralumin half-cylinder and the hydrophone needle is moved in a 3D domain.

any point in front of the cavity. This ''synthetic time-reversal'' technique is validated by experimental results.

II. PRINCIPLE OF TIME REVERSAL WITH A CAVITY

The study is restricted to the scalar representation of the acoustic field. The source is defined by the source function $s(A,t)$. In the water, the signal $r(B,t)$ on the receiver is given by the Rayleigh's integral

$$r(B,t)=\int_{t',A}\int g(B,t|A,t')s(A,t')dA\,dt'. \tag{1}$$

In Eq. (1) g is the Green's function between a point source in A to an observation point located in B. The Green's function is not invariant by spatial translation, because the medium is finite and inhomogeneous. Then, Eq. (1) cannot be reduced to a simple spatial convolution product. However, the invariance is verified in the temporal dimension and the Green's function can be rewritten

$$g(B,t|A,t')=g(B,t-t'|A). \tag{2}$$

If we assume a uniform temporal excitation over the whole source (piston source), the function $s(A,t)$ is expressed as: $s(A,t)=s(A)\cdot e(t)$, with $s(A)$ constant. In this case, Eq. (1) reduces to

$$r(B,t)=\int_{t',A}\int g(B,t-t'|A)s(A)dA\,e(t')dt'. \tag{3}$$

Thus, the signal corresponding to a position B_o can be written in the form of a temporal convolution product

$$r(B_o,t)=h(A,B_o,t)\otimes_t e(t), \tag{4}$$

where h (i.e., the impulse response) is the double integral of the Green's function over the spatial dimensions. By taking into account the spatial reciprocity, if this received signal is time reversed and re-emitted, the receiver observes in an arbitrary position B

$$r_{\mathrm{TR}}(B,t)=h(A,B,t)\otimes_t h(A,B_o,-t)\otimes_t e(-t). \tag{5}$$

In an ideal time-reversal experiment, the emitted field is perfectly reconstructed and one would obtain

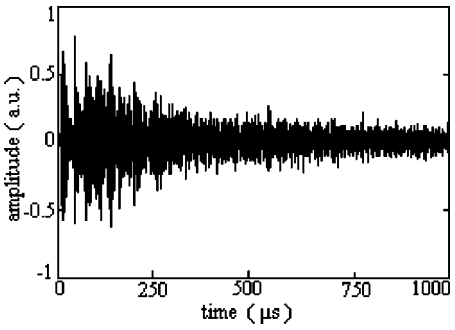


FIG. 2. Amplitude of the beginning of an impulse response given by the hydrophone needle.

$$r_{\mathrm{TR}}(B,t)=\delta(B-B_o)\cdot\delta(t-t_0)\otimes_t e(-t), \tag{6}$$

where t_0 and $\delta(B-B_o)\cdot\delta(t-t_o)$ stand for the recompression time and the Dirac distribution, respectively. When the impulse responses $h(A,B,t)$ aren't reduced to a simple Dirac distributions, the condition to come close to such a field is that the impulse responses $h(A,B,t)$ are decorrelated. In real situations, we expect to obtain as few time-correlated signals as possible using a cavity, owning both ergodic and, from an acoustic ray point of view, chaotic properties. Our experimental setup, Fig. 1, uses a duralumin half-cylinder cavity, of radius 50 mm and height 100 mm. A 1-MHz contact transducer is stuck on its back surface and a hydrophone needle is used to record the radiated field in the water. Both sensors are connected to a one-channel time-reversal electronic driven by a computer. In such cavities, when a source sends a short pulse (a 1.5-cycle of 1-MHz sine wave), the scattered signal spreads over 1000 μs, i.e., 500 times the initial pulse duration. A typical impulse response from the contact transducer to the hydrophone needle is plotted in Fig. 2. It looks like a signal of a diffused field with a decreasing exponential envelope due to the absorption and radiation. The decorrelation property of such signal assures a time-reversal process very close to the ideal result of Eq. (6).

Such an impulse response is recorded at 100 mm from the cavity, time reversed and re-emitted with the contact transducer. The radiated field is scanned using the hydrophone needle in the focal plane, with a spatial sampling grid of a half wavelength (0.75 mm). The normalized amplitude of the field is plotted in Fig. 3(a) at three different times. At the recompression time (chosen as the origin of time), a recompression peak emerges on the location where the impulse response was recorded: it is the wave focusing back to its source. At 2 μs before and after the recompression time, one can observe an average level at -15 dB. As it is shown in Fig. 3(b), the spatial dimensions of the recompression peak are 20 times smaller than the directively of the transducer alone in the water. From now on, the diffracting aperture is the front face of the cavity. As the directivity pattern is improved, the ultrasonic focusing can be effective anywhere in a 3D domain much larger than the ultrasonic beam of the transducer in the water. Multifocusing is also possible with this technique.

To illustrate the real-time 3D focusing property, the following experiment is done: a moving point source is emitting pulses every 40 ms (real time) in water in a domain where a

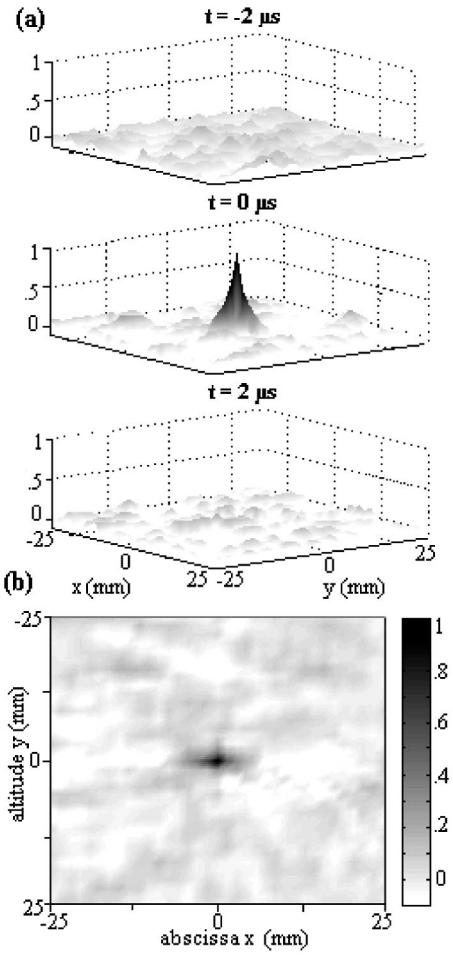


FIG. 3. (a) Spatio-temporal ultrasonic field and (b) projection of the maximum of the field on the focal plane, during an experimental time-reversal focusing at 100 mm from the half-cylinder. The origin of time is the recompression time.

whole set of impulse responses is first collected in a ''library.'' Then, each impulse response recorded on the contact transducer is correlated with those of the library. This computation is a special beamforming in receive mode (matched filter), where complicated Green's functions are used. The correlation maxima are shown in Fig. 4, in gray-scale images extracted from a movie. Each image characterizes the ultrasonic field and indicates the point source location at different

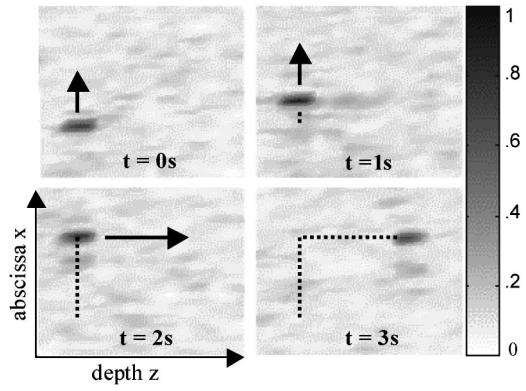


FIG. 4. Four images extracted from a movie showing the movement of a pulsing source in the water. The observation area is 5×5 cm.

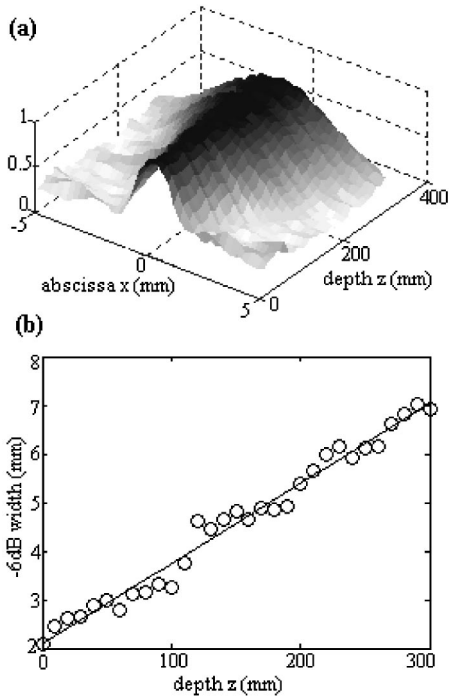


FIG. 5. (a) Profiles of focal zone for focusing at different depth. (b) The -6-dB width of focal spot as a function of depth.

times. Using the reciprocity property these images of Fig. 4 can also be interpreted in transmit mode as the focusing field emitted by the cavity every 40 ms. The minimum time between each focusing process depends on the decreasing speed of the impulse response, which is given by both dissipation and radiation properties of the cavity. In this experiment this minimum time between focusing process is estimated at 500 μs.

III. DIFFRACTION

The aim of this section is to show how the diffraction laws can emerge from the apparent disorder of the ultrasonic radiated field. Figure 5 presents 30 profiles of the field amplitude along a line of the focal plane (a) and their -6-dB width (b). Each of these directivity patterns is obtained by focusing ultrasounds at distances ranging from 10 to 400 mm. The focusing width Δx variations with depth correctly verify the diffraction theoretical law $\Delta x = \lambda(F/D)$ (λ is the wavelength, F the focusing depth, and D the aperture width). Therefore, the fit of the experimental measurement of Δx as a function of z allows one to deduce apparent diameter aperture. It is found to be equal to $D = 95$ mm, which is almost the lateral dimension of the front face of the cavity (100 mm).

This result indicates that all the conversion of elastic energy inside the cavity into acoustic energy in the focusing domain is carried out on the whole surface. This confirms the assumption about the ergodic property of the present cavity. A deeper mathematical analysis of the diffraction laws applied on the ultrasonic field will lead us, in the next section, to define a simpler time-reversal protocol for our experiment: ''the synthetic time-reversal technique.''

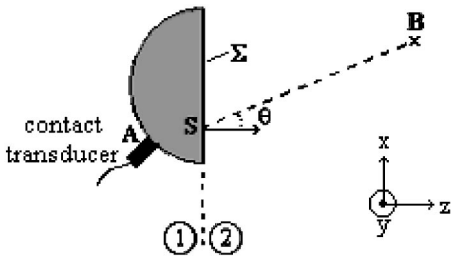


FIG. 6. Sketch of the setup and notations.

As it is shown in Fig. 6, the vertical plane which contains the front face of the cavity defines two domains (1) and (2). The first one contains the cavity with the contact transducer and the second one is the water where the hydrophone can record the pressure at any point B. We call Σ the radiation surface of the cavity and S a point of this aperture. Owing to the Rayleigh's integral with Dirichlet boundary conditions, the pressure on point B is deduced from the field measured on the radiation aperture

$$p(B,t)=\frac{jk}{2\pi}\int_{\Sigma}\left[\frac{\partial p(S,t)}{\partial t}\underset{t}{\otimes}G^{L}(S,B,t)\right]\cdot\cos\theta\cdot dS. \quad (7)$$

In Eq. (7), $p(S,t)$ represents the pressure field on the radiation aperture and $G^{L}(S,B,t)$ the Green's function defined by

$$G^{L}(S,B,t)=\frac{\delta(t-|SB|/c)}{|SB|}. \quad (8)$$

If a short pulse $e(A,t)$ is emitted by the transducer set on point A, the pressure at a point S of the cavity's front face is described by the following expression:

$$p(S,t)=e(A,t)\underset{t}{\otimes}h(A,S,t), \quad (9)$$

where $h(A,S,t)$ is the impulse response between the transducer and a point S of the radiation aperture. By taking into account Eq. (9), the pressure on the point B can be written as the following:

$$p(B,t)=e(A,t)\underset{t}{\otimes}\frac{jk}{2\pi}\int_{\Sigma}\frac{\partial h(A,S,t)}{\partial t}\underset{t}{\otimes}G^{L}(S,B,t)\cdot\cos\theta\cdot dS. \quad (10)$$

The term on the right-hand side of the convolution product can be identified to the impulse response $h(A,B,t)$ which connects a pulse injected from the transducer at point A, to the pressure field at point B

$$h(A,B,t)=\frac{jk}{2\pi}\int_{\Sigma}\frac{\partial h(A,S,t)}{\partial t}\underset{t}{\otimes}G^{L}(S,B,t)\cdot\cos\theta\cdot dS. \quad (11)$$

Because of the spatial reciprocity this impulse response, $h(A,B,t)$ is equal to $h(B,A,t)$ and the time-reversed signal $p(B,-t)$ can be injected in the cavity by the transducer: the generated field propagates as if it was traveling backwards in time. Thus, the pressure on point B during the time-reversal experiment is written as a convolution between the impulse responses

$$p_{\mathrm{TR}}(B,t)=e(A,-t)\underset{t}{\otimes}h(A,B,-t)\underset{t}{\otimes}h(A,B,t). \quad (12)$$

Using the expression of the impulse response $h(A,B,t)$, the pressure on B can be deduced from Eq. (11) and Eq. (12)

$$p_{\mathrm{TR}}(B,t)=\frac{jk}{2\pi}\int_{\Sigma}\frac{\partial p(S,-t)}{\partial t}\underset{t}{\otimes}G^{L}(S,B,-t)\cdot\cos\theta\cdot dS\underset{t}{\otimes}h(A,B,t). \quad (13)$$

In this expression, the term in front of the impulse response $h(A,B,t)$ stands for the emission of a classical time-reversal experiment. Equation (13) shows that this emission could be constructed based on the knowledge of the pressure field on the cavity's face, $p(S,t)$.

In real terms the field on the cavity's front face is scanned with a hydrophone needle, according to a sampling grid with a spatial step equal to the wavelength (correlation length in a chaotic cavity), in order to measure the impulse responses $p(S,t)$. By taking into account this spatial sampling, the emission signal deduced from Eq. (13) can be written as

$$e(t)=e(A,-t)\underset{t}{\otimes}h(A,B,-t)=\frac{jk}{2\pi}\sum_{i}\frac{\partial p^{\exp}(S_i,-t)}{\partial t}\underset{t}{\otimes}G^{L}(S_i,B,-t)\cdot\cos\theta_i. \quad (14)$$

Each experimental impulse response $p^{\exp}(S_i,t)$ is time reversed, once time derived, and then convoluted with a computed free-space Green's function. The ensemble of these Green's functions, which corresponds to a point source in the water on B, obeys a spherical delay law and a decrease as $1/|SB|$. Finally, a summation over the whole set of signals gives the exact emitted signal able to focus on an arbitrary focal point (point B) defined by the free-space Green's function.

In other words, this ''synthetic time-reversal process'' permits one to focus ultrasounds anywhere in a 3D domain, with the time-reversal efficiency, just by recording the impulse responses on a 2D surface (the surface of the cavity) instead of a 3D volume (the water in front of the cavity).

First, the emission of the synthetic time reversal can be compared with the emission of the classical time-reversal process. Figure 7 shows 20 μs of both signals: their correlation coefficient reaches 0.89. This good result is not perfect for two reasons. The first one is the practical sampling of the scan of the front face, which infers a loss of information, because we have chosen the spatial step equal to the wavelength. The second one is that the cavity rests on a support, which can slightly radiate ultrasounds. This radiation is not taken into account in the synthetic time-reversal process, but is present in a real-time reversal experiment.

Figure 8 allows one to compare the two profiles of the focal peak at the point B: the gray dotted one is obtained by the classical time-reversal process and the black one by the synthetic time-reversal process. The sidelobe amplitude is comparable for the two experiments and remain on average

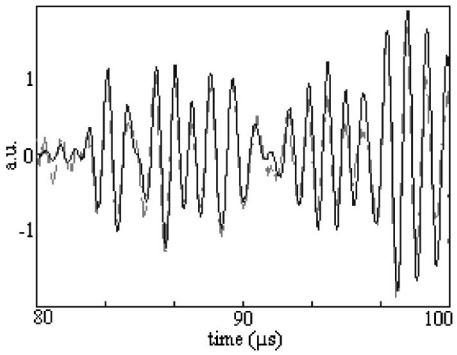


FIG. 7. Details (20 μs) of the emissions of the real time-reversal process (gray dotted line) and the "synthetic time-reversal process" (black line). The correlation coefficient between the entire emissions is equal to 0.89.

at −15 dB. On the other hand, the −6-dB width of the focal spots are both equal to 2 mm (1.3 wavelength).

Thus, the knowledge of the impulse responses of the 2D cavity's aperture permits computing of all the impulse responses of a volume in front of the cavity. It means, in other words, that the synthetic time-reversal technique can focus ultrasounds in a 3D domain with the same quality as a real-time reversal experiment, without the disadvantage of collecting a whole set of impulse responses on this domain. The acquisition protocol of the library of reference signals is thus simplified.

IV. DISCUSSION

In this last part, we will try to go further to understand the ultrasonic field radiated by the cavity during a time-reversal experiment. A measurement of the field on the front face of the cavity during reemission of the time-reversed impulse response of a point located at 10 cm gives the experimental B-scan of Fig. 9. A focusing wavefront can be seen in the spatio-temporal sidelobes.

In order to describe the field radiated by the cavity's front face, Eq. (13) can be developed to give the following expression:

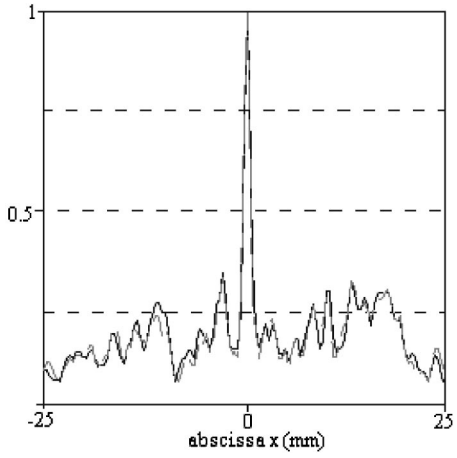


FIG. 8. Profiles of focused field's maximum amplitude, obtained by synthetic time-reversal process (black line), or by real-time reversal process (gray dotted line).

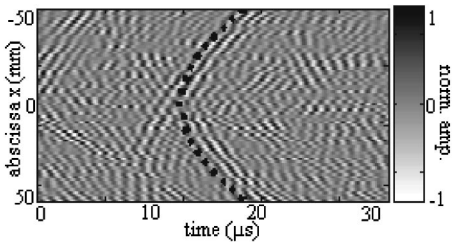


FIG. 9. B-scan of the radiating field on the cavity's face, after the emission of the synthetic time-reversal signal. The dotted line shows the perfect spherical delay law.

$$p(B,t) = -\frac{k^2}{4\pi^2}\cdot e(A,-t)\underset{t}{\otimes}\int_{\Sigma}\frac{\partial h(A,S,-t)}{\partial t} \underset{t}{\otimes} G^L(S,B,-t)\cdot\cos\theta\cdot dS, \underset{t}{\otimes}\int_{\Sigma}\frac{\partial h(A,S',t)}{\partial t}\underset{t}{\otimes} G^L(S',B,t)\cdot\cos\theta'\cdot dS'. \tag{15}$$

In this expression, S and S' are two independent variables. Equation (15) can be rewritten

$$p(B,t) = -\frac{k^2}{4\pi^2}\cdot e(A,-t)\underset{t}{\otimes}\int_{\Sigma}\int_{\Sigma}\frac{\partial h(A,S,-t)}{\partial t} \underset{t}{\otimes} G^L(S,B,-t)\cdot\cos\theta\cdot dS, \underset{t}{\otimes}\frac{\partial h(A,S',t)}{\partial t}\underset{t}{\otimes} G^L(S',B,t)\cdot\cos\theta'\cdot dS'. \tag{16}$$

By comparing the above equation with the Rayleigh's integral Eq. (7), the pressure on point S' of the cavity's front face can be identified

$$\frac{\partial p(S',t)}{\partial t} = \frac{jk}{2\pi}\cdot e(A,-t)\underset{t}{\otimes}\int_{\Sigma}\frac{\partial h(A,S,-t)}{\partial t} \underset{t}{\otimes} G^L(S,B,-t)\cdot\cos\theta\cdot dS\underset{t}{\otimes}\frac{\partial h(A,S',t)}{\partial t}. \tag{17}$$

By integrating as a function of time, we obtain the expression that we were looking for

$$p(S',t) = \frac{jk}{2\pi}\cdot e(A,-t)\underset{t}{\otimes}\int_{\Sigma}\frac{\partial h(A,S,-t)}{\partial t} \underset{t}{\otimes} G^L(S,B,-t)\cdot\cos\theta\cdot dS\underset{t}{\otimes} h(A,S',t). \tag{18}$$

Equation (18) gives the ultrasonic field on the front face of the cavity during a time-reversal experiment. If $S=S'$, the autocorrelation $h(A,S,-t)\otimes_t h(A,S',t)$ describes the temporal recompression peak. According to the time-reversal theory, this recompression peak gives birth to a divergent wave, such that the point S can be considered as a secondary

source. If S describes the whole surface of the cavity, these secondary sources emit a temporal recompression peak according to the spherical delay law imposed by the free-space Green's function, $G^L(S,B,-t)=\delta(-t-[|SB|/c])/|SB|$. Thus, Eq. (18) can be interpreted as a consequence of the Huygens–Fresnel principle since a spherical wavefront is created from secondary sources: the front face of the cavity works as a 2D transducers array.

The signal-to-noise ratio can be now discussed. The ultrasonic sidelobes are linked to the correlation $h(A,S,-t) \otimes_t h(A,S',t)$ which describes the temporal recompression if $S'=S$, but also the sidelobes if $S' \neq S$. For each secondary source on the surface of the cavity, the sidelobes of all the other sources is added to the temporal recompression. This qualitative consideration explains why the spherical wavefront is not clearly apparent at first sight.

From a quantitative point of view, the signal-to-noise ratio can be shown to be the same inside (on the surface) and outside (in front) of an ergodic cavity. Let P be the peak amplitude and N the average noise amplitude of a time-reversal experiment on a point of the surface of the cavity. In this case, the signal-to-noise ratio is P/N. When this particular point is one of the M secondary sources of a spherical wavefront, the signal-to-noise ratio must take into account the additive sidelobes due to the $M-1$ other sources. Since the noise increases as $\sqrt{M}$ for decorrelated secondary sources, the decreasing law of this signal-to-noise ratio becomes $P/(\sqrt{M}\cdot N)$. At last, a coherent summation of the recompression peak is achieved during the propagation in the water. Thus, the peak increases as M whereas the noise still evolves as $\sqrt{M}$. The resulting signal-to-noise ratio on the focal point $(M\cdot P)/(\sqrt{M}\cdot\sqrt{M}\cdot N)=P/N$ is equal to the ratio on the surface. Indeed, the experimental peak-to-noise ratio on the cavity's front face is found to be 20.3 dB. On the focal point, this ratio is 21.7 dB, which validates the previous arguments. It is important to note that this result is valid on average (ergodic cavity).

V. CONCLUSION

We have shown the feasibility to focus ultrasounds in real time, on a 3D domain, with a single transducer coupled to a solid structure and the time-reversal process. The focusing quality depends on the ability of the solid to ''mix'' the ultrasonic field emitted by the transducer so that each point has a unique signature. This condition is achieved using a cavity owning both ergodic and chaotic properties.

A careful study of the spatial focal spot in a water tank has shown that this focusing system agrees with the diffraction laws. Our understanding of the diffraction physics leads to simplify the experimental procedure through the synthetic time-reversal technique: theory and experimental results show that the knowledge of the impulse responses on the cavity's front face allows one to focus ultrasonic field on any point in front of the cavity.

Our future works will be to increase the peak-to-noise ratio of the spherical wavefront formed on the surface of the cavity. By improving the noise level of the focusing on emission and reception, our hope is to obtain a real-time echographic imaging device with a single transducer.

[1] M. Fink, ''Time reversed acoustics,'' Phys. Today **50**, 34–40 (1997).
[2] C. Draeger and M. Fink, ''One channel time reversal of elastic waves in a 2D silicon cavity,'' Phys. Rev. B **79**(3), 407–410 (1997).
[3] C. Draeger and M. Fink, ''One channel time reversal in chaotic cavities: theoretical limits,'' J. Acoust. Soc. Am. **105**(2), 611–617 (1999).
[4] C. Draeger and M. Fink, ''One channel time reversal in chaotic cavities: Experimental results,'' J. Acoust. Soc. Am. **105**(2), 618–625 (1999).
[5] J. de Rosny, A. Tourin, and M. Fink, ''Coherent backscattering of an elastic wave in a chaotic cavity,'' Phys. Rev. Lett. **84**(8), 1693–1695 (2000).
[6] A. Derode, A. Tourin, and M. Fink, ''Random multiple scattering of ultrasound. II. Is the time reversal a self-averaging process?,'' Phys. Rev. E **64**, 036606–1 036606–13 (2001).
[7] J. A. Root and P. H. Rogers, ''Performance of an underwater acoustic volume array using time reversal focusing,'' J. Acoust. Soc. Am. **112**(5), 1869–1878 (2002).

Surface acoustic wave modulation on a partially closed fatigue crack

J.-Y. Kim, V. A. Yakovlev, and S. I. Rokhlin[a)]
Nondestructive Evaluation Program, The Ohio State University, Edison Joining Technology Center, 1248 Arthur E. Adams Dr., Columbus, Ohio 43221

(Received 29 September 2003; revised 12 February 2004; accepted 15 February 2004)

A quantitative study of the low-frequency parametric modulation of a pulsed surface acoustic wave (SAW) by a partially closed fatigue crack is described. *In situ* ultrasonic measurements were performed during a fatigue test for different crack lengths and static opening loads. The crack is initiated in the plastic-yielding zone induced by a surface cavity, and clamped due to the constraint of the surrounding elastic medium. Small periodic loading, superimposed on a static crack-opening load, changes the open crack segment length and/or the crack interfacial condition producing nonlinear modulation of the reflected ultrasonic pulses. The modulation spectrum is related quantitatively to the crack length and to the crack opening–closure behavior. It is demonstrated that the application of a small static crack-opening load with the modulation load could considerably enhance crack detectability. The increase of the second modulation harmonic is pronounced when the crack is nearly closed and when it is nearly open. Also, it is observed that the maximum modulation occurs at different static opening loads depending on the crack length relative to the plastic-yielding zone size. A low-frequency scattering model is presented based on the mechanism of modulation of the open/close segment length of the partially opened crack. The modeling results compare favorably with experiment. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1695012]

PACS numbers: 43.25.Fe, 43.25.De, 43.25.Zx [4HB] Pages: 1961–1972

I. INTRODUCTION AND PROBLEM STATEMENT

The success of conventional linear ultrasonic techniques for detecting small fatigue cracks is often limited by ultrasonic grain noise, existence of other volumetric inhomogeneities, and crack closure. Nonlinear acoustic[1–5] and thermal[6] methods have been found to be promising for material diagnostics and damage detection in complex media since it has been established that imperfect interfaces associated with damage generate nonlinearity anomalously higher than in bulk solids. Higher harmonic[1,2,4] and subharmonic generation and chaotic behavior[6] have been observed from small cracks by nonlinear acoustic methods. The related mechanisms were Hertzian contact at microcrack surface asperities[8] and/or clapping motions of the crack interfaces, depending on the amplitude of the incident ultrasonic wave.[2,6] Recently, Zaitsev *et al.*[9] suggested a thermoelastic mechanism which explains the nonlinear effects by heating and cooling of the crack perimeter to interpret the logarithmic slow dynamic and relaxation[10] during interaction of elastic vibrations with material damage. Moussatov *et al.*[11] observed experimentally a self-induced nonlinear hysteresis behavior for the bulk sinusoidal acoustic wave interacting with a system of cracks inside a glass plate.

A somewhat different approach in nonlinear ultrasonic diagnosis is modulation of an ultrasonic wave on an imperfect interface or a crack by an external dynamic load with a frequency much lower than the probing ultrasonic wave. The modulation can be induced thermally[12,13] using laser pulse irradiation or by different mechanical means.[14–17] In medical ultrasonic diagnostics, a parametric method called elastography is used to distinguish different pathological objects by applying static or low-frequency dynamic loading.[18,19]

Different mechanisms may be responsible for the nonlinear parametric modulation of acoustic waves interacting with a crack, depending on the total stress acting on the crack. The total stress results from an external load (including static and/or dynamic modulation loads) and the intrinsic residual stress.[20,21] For example, when the modulation stress is much higher than the intrinsic closure stress, clapping motion occurs and produces nonlinear modulation.[1,2,6] When the modulation stress and the crack closure stress are comparable, the crack opens gradually following the modulation load. This is accompanied by change of the open crack segment length[22] and by change of the finite stiffness contact[12] at the crack tips with increased level of the crack tip modulation signal.[17] Often both mechanisms may exist. Since the ultrasonic wave reflectivity from the fatigue crack depends significantly on the condition on the crack interface,[20,21,23] the modulation response of a crack is also strongly influenced by the level of crack closure. For example, when a crack is fully open it behaves like a volumetric flaw and thus the output signal is weakly modulated. When the crack is tightly closed the modulation of the output signal is also small. Therefore, to apply the modulation technique as a diagnostic tool one needs a quantitative understanding of the crack state as it is related to the parametric modulation of ultrasonic waves. In most prior studies, the actual crack closure stress was unknown and thus the related modulation mechanisms were difficult to identify.

a)Electronic mail: rokhlin.2@osu.edu

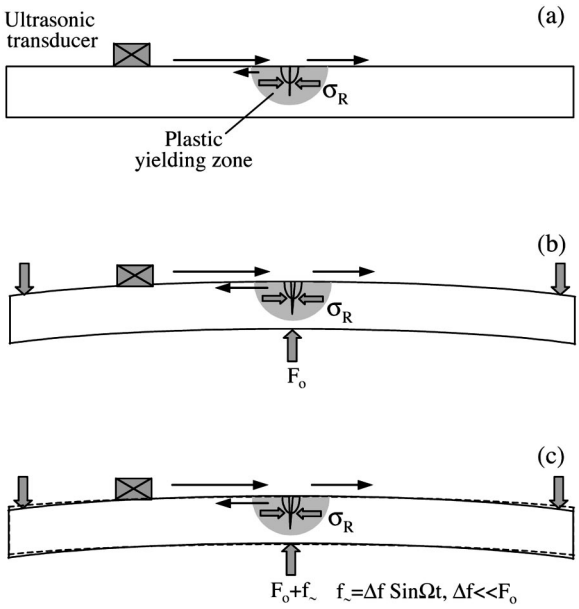


FIG. 1. Schematic showing the surface acoustic wave measurement for detecting a surface crack initiated in the plastic-yielding zone of a surface cavity. (a) Due to crack closure by compressive residual stress (σ_R), the crack is transparent to the ultrasonic wave and only the cavity reflection is measured. (b) A static external load F_o is applied to open the crack during the ultrasonic measurement. (c) A periodic modulation load ($f_\sim$) with small amplitude is applied while the reflection pulses are taken synchronously with the periodic vibration. *(In this figure, due to practical relevance, bending quasistatic forces are applied to the structure as illustration. In the experiments described in this paper, the load is applied in tension.)*

Early detection of fatigue crack initiation from corrosion damage is extremely important for determination of remaining fatigue life.[24,25] Thus, improvement of small crack detectability is very important. In this paper, we investigate the effect of low-frequency modulation on pulsed ultrasonic surface wave interaction with a fatigue crack initiated in the plastic-yielding zone of a surface cavity. It expands a recently published letter[22] by detailed discussion of experiments and a low-frequency scattering model explaining the experimental data. The ultrasonic measurements are performed *in situ* during a fatigue test acquiring the reflected surface wave pulse at different levels of closure stress. The different levels of partial opening of a closed crack are formed by applying a variable static opening load. It is shown that the modulation response depends on both the applied static and dynamic modulation loads. The influence of crack closure on the modulation spectrum for different crack lengths is measured and physically interpreted with the mechanism of change of the open crack segment length. The low-frequency scattering model is applied to predict the modulation and higher harmonic amplitudes.

To illustrate the general concept of ultrasonic reflectivity from a partially closed fatigue crack, Fig. 1(a) shows a surface crack which is initiated in the plastic-yielding zone of a surface cavity (pit). The crack is under compressive stress and thus partially closed.[20–23] The backscattering signal measured by an ultrasonic transducer is formed by the reflection from the cavity and by the contribution from the crack. The surfaces of a fresh fatigue crack almost conform, which makes the crack nearly transparent to the ultrasonic wave, when it is closed. Thus, detection of such a crack is very difficult. In this case, one may attempt to open the crack applying external force during measurement as shown in Fig. 1(b) (bending forces are applied for illustration only). However, the signal reflected from the cavity may still dominate the total reflected signal and a considerably higher force may have to be applied to open the crack for its detection. This may be impractical for a large structure. A modulation technique is illustrated in Fig. 1(c), in which a low-frequency vibration is applied to the sample; it is also illustrated schematically in Fig. 2. A sequence of ultrasonic pulses is injected into the sample with a simultaneous application of low-frequency vibration. The sample properties (crack closure condition) change synchronously with the modulation vibration affecting the amplitude of the ultrasonic pulses interacting with the medium (for example, with the crack). The modulated ultrasonic pulses are demodulated to extract the amplitude change. Imperfect interfaces or cracks usually generate much higher modulation than a volumetric flaw does.[12] In this way, the effect of the cavity, which is unchanged and forms a major part of the ultrasonic signal amplitude, is virtually removed and thus the change of the pulse amplitudes due only to the crack is obtained and crack detection is significantly enhanced. The efficiency of the modulation which is indicated by the change of pulse amplitude divided by the average pulse amplitude (modulation index) depends largely on the mechanism of ultrasonic modulation from the crack.

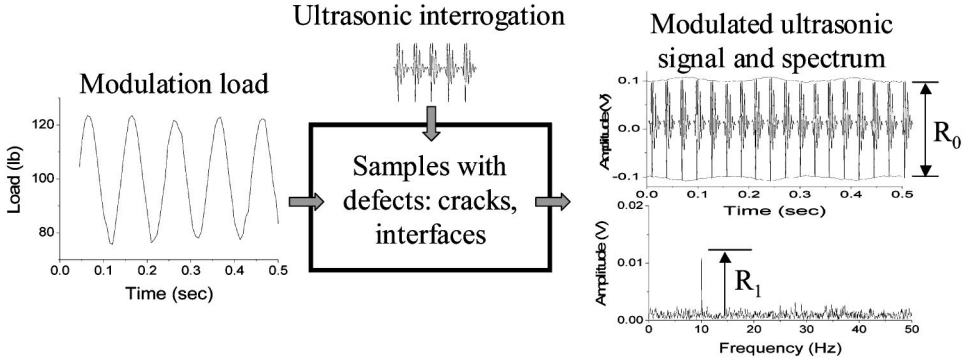


FIG. 2. Modulation technique using pulse sequence with the harmonic modulation load. The modulation index is defined to be the ratio between the modulated signal amplitude and the average pulse amplitude (R_1/R_0).

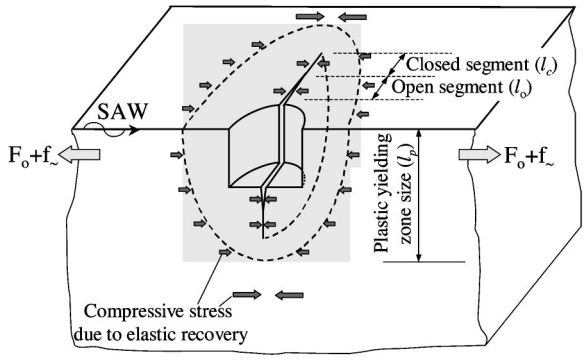


FIG. 3. Schematic showing fatigue crack initiated in plastic-yielding zone. Crack is under compressive residual stress due to elastic recovery of the surrounding medium. Lengths of open (l_o) and closed (l_c) crack segments and plastic-yielding zone size (l_p) are indicated.

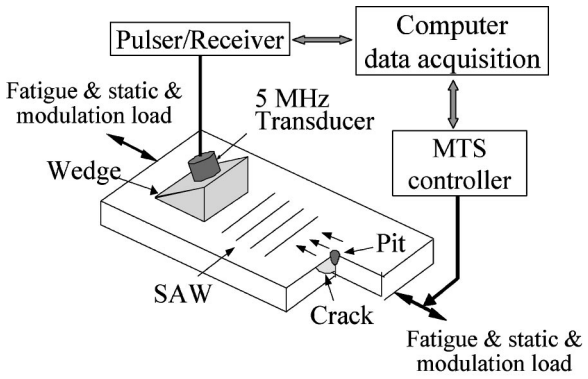


FIG. 4. Setup for *in situ* ultrasonic experiment during fatigue tests.

II. EXPERIMENT

A. Sample and fatigue test

The fatigue sample was machined from 1.6-mm-thick Al 2024-T3 alloy plate with 340 MPa the yield stress, 483 MPa ultimate tensile stress, and 17.5% elongation. For crack initiation a controlled-size small pit with depth of 250 μm and diameter of 230 μm was produced by an electrical discharge machine (EDM) in the center of the sample. The frequency of the fatigue load was 10 Hz; the stress ratio ($\sigma_{\min}/\sigma_{\max}$) was 0.1, and the stress range $\Delta\sigma$ was 231 MPa. The maximum fatigue stress level was 78% of the yield stress so that the high stress concentration ($k_t \approx 3.45$) at the pit corners leads to formation of a plastic-yielding zone.[20,23] At load reduction the enlarged yielded volume does not recover to its original size, and thus is constrained by the compressive stress due to elastic recovery of the surrounding medium as shown in Fig. 3. Therefore, the crack initiated in the plastic-yielding zone is closed by the compressive closure stress.[20,21] The fatigue life of the samples was about 150 kcycles. At this number of cycles the crack has almost transited through the sample thickness and the sample is near failure.

B. *In situ* quasistatic and modulation ultrasonic measurements

Ultrasonic modulation measurements were performed for different crack lengths relative to the plastic-yielding zone. This is achieved by *in situ* ultrasonic measurements during a fatigue test with predetermined intervals at different stages of crack initiation and evolution and at different levels of crack opening load. The experimental setup includes a 5-MHz ultrasonic surface wave transducer, a pulser/receiver, a mechanical testing system (MTS), and an ultrasonic data acquisition as shown in Fig. 4 For an *in situ* measurement, the transducer assembly is clamped on the sample so that the ultrasonic signals are collected during fatigue and modulation load cycling. An ultrasonic couplant was applied between the ultrasonic wedge and the sample. Software developed in-house was used for fatigue and modulation control and for synchronized real-time ultrasonic data acquisition. A software control 12-bit, 125-MHz digitizing computer board was used for data acquisition. The data acquisition and the ultrasonic pulser/receiver are triggered by a software control counter at predetermined numbers of fatigue cycles (crack depths) and fatigue and modulation loads. The ultrasonic measurements were performed in the pulse–echo mode with a repetition frequency 300 Hz.

Load profiles versus time (number of cycles) with indicated times of ultrasonic measurements are schematically shown in Fig. 5. Loading sequence and ultrasonic data acquisition have been preprogramed for automatic execution. The fatigue experiment is divided in repetitive segments. Each segment consists of three parts [Fig. 5(a)]: (1) Part I is cycling fatigue, during which no ultrasonic signals are taken (this is to avoid an enormous quantity of collected ultrasonic data). (2) Part II consists of a single loading–unloading fa-

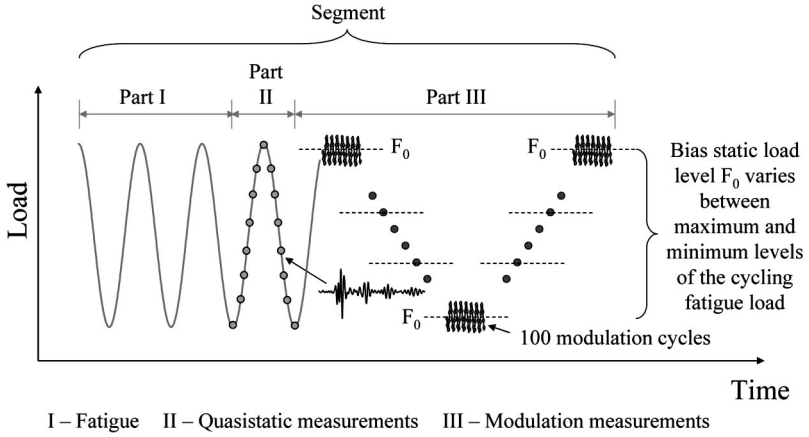


FIG. 5. The entire fatigue experiment has been divided into identical segments where ultrasonic data have been collected. One of the repetitive segments is shown schematically. The segment consists of three parts: part I is the cycling fatigue part (no ultrasonic signals are taken); part II consists of a single loading–unloading fatigue cycle where excitation and data acquisition of ultrasonic surface wave signatures are performed. Ultrasonic data collection occurs a total of 30 times per cycle (shown by gray circles). Modulation measurements at different levels of static bias load F_o are performed in part III of the segment. Thirty ultrasonic reflection signals per modulation cycle were acquired synchronously for over 100 cycles. After part III the fatigue cycling is resumed, starting part I of the next identical segment.

tigue cycle where excitation and data acquisition of ultrasonic surface wave signatures are performed at different load levels in the cycle (shown by gray circles). Ultrasonic data collection occurs 30 times per cycle in these series of experiments. Since an ultrasonic event occurs during a very short time period, the fatigue load during the event can be considered quasistatic. (3) Immediately after the quasistatic ultrasonic measurements follows Part III of the segment, where the modulation measurements are performed by superimposing the modulation vibration ($f_{\sim} = \Delta f_{\sim} \cos \Omega t$) at different levels of static bias load F_o as shown in Fig. 5(b). The amplitude of the modulation load ($\Delta f_{\sim}$) was as low as 0.17 kN (17.4 MPa) at 10 Hz. The static load (F_o) varies with 0.2 kN (20.5 MPa) steps from 0.31 kN (30.8 MPa) to 2.36 kN (242 MPa) (between maximum and minimum levels of the cycling fatigue load). Thirty ultrasonic reflection signals per modulation cycle were acquired synchronously for over 100 cycles. After performing the ultrasonic measurements at different levels of static load with modulation, the fatigue cycling resumed, starting the next identical loading/data-acquisition segment.

The collected series of ultrasonic signatures represents pulse amplitude-modulated sequences at different static loads (crack closures) and different numbers of cycles (crack lengths). The acquired pulses were demodulated by low-pass filtering and cross correlated with the modulation load for signal-to-noise ratio improvement. Measurements have been repeated on several samples demonstrating consistent results.

In our experiment both reflected and transmitted waves are measured; however, only reflected wave data are reported in this paper.

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Quasistatic measurements

A typical reflected ultrasonic signal from a pit without a crack is shown in Fig. 6. The surface acoustic wave reflection from a pit is composed of waves reflecting from different corners of the pit, followed by the plate bottom reflections of a mode-converted shear wave, Fig. 6(a). The first group consists of the specular reflection from the front of the pit (1), the corner creeping wave on the edge of the pit (2) (see Ref. 24 for the creeping wave discussion), and the reflection signal going down around the pit and reflected from the pit opposite corner (3). Although the signals in the first group are not well separated, the time delay between the signals can be calculated from the frequency spectrum of the gated signal.[21]

The second signal is the plate bottom reflection [4 in Fig. 6(a)] of the shear wave. The mode-converted shear wave (4) is launched on the bottom of the pit and propagates down toward the bottom of the plate where it is reflected, mode converted back to the surface wave at the bottom of the pit, and returned to and received by the surface wave transducer. This signal is separated from the pit-reflected signals in the time domain, Fig. 6(b). As one can see in Fig. 6(b), the plate bottom reflection signal is narrower and better defined than that of the pit reflection. The bottom reflection provides useful information on crack closure behavior and one can get the depth of the surface discontinuity by analyzing the time delay of the bottom reflection signal.

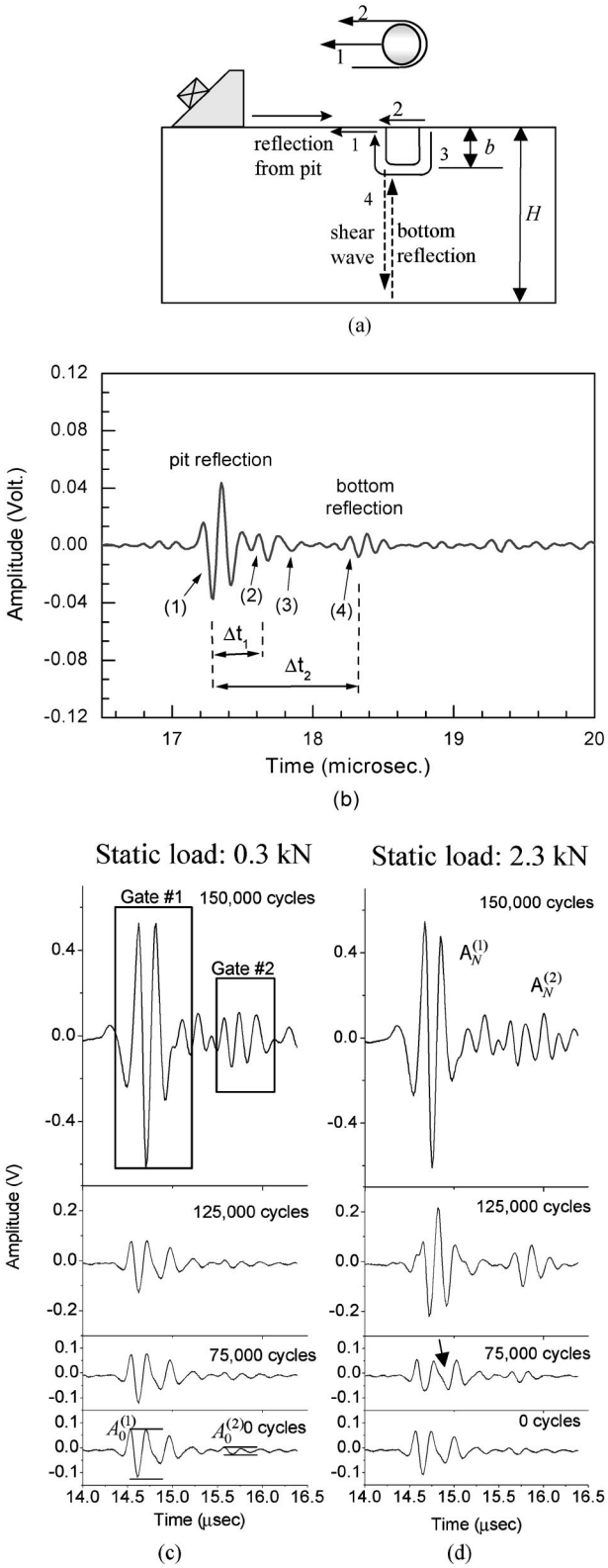


FIG. 6. Time-domain surface wave reflection signals (a,b) from a pit without crack. (a) Different paths of reflected waves; (b) Reflected signals with different paths are marked. (c,d) Ultrasonic signals acquired at different fatigue cycles at two different static loads: (c) 0.3 kN; (d) 2.3 kN. The fatigue life of the sample was 152,531 cycles.

Typical reflected ultrasonic signatures from a pit with an emanating fatigue crack are shown in Figs. 6(c) and (d) at

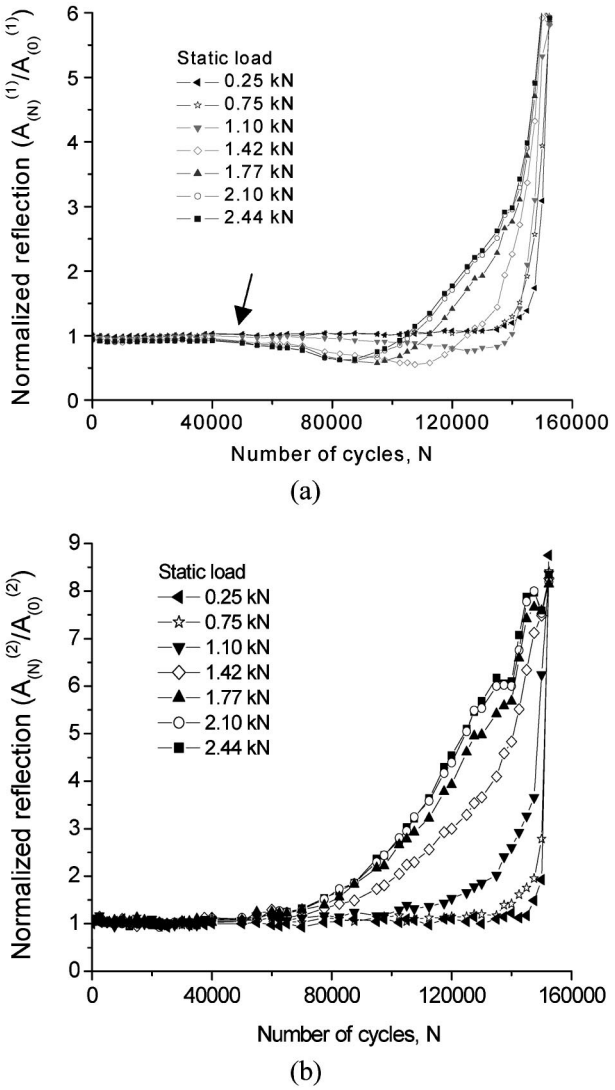


FIG. 7. Change of normalized reflection signal amplitudes versus number of fatigue cycles for different load levels. (a) First reflection; (b) Plate bottom reflection.

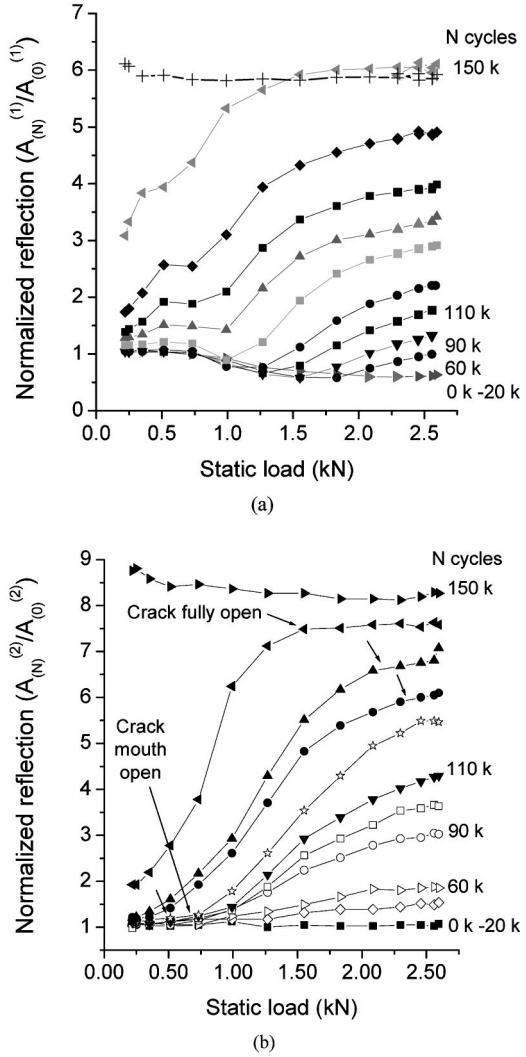


FIG. 8. Normalized reflection versus load at different number of cycles. (a) First reflection; (b) Plate bottom reflection.

different numbers of cycles for two load levels. The surface acoustic wave reflections are composed of waves reflecting from the pit with crack ($A_N^{(1)}$), followed by the plate bottom reflection ($A_N^{(2)}$) of a mode-converted shear wave, where the subscript N represents the number of cycles. As a result of crack initiation and growth during the fatigue test the amplitudes of both the first ($A^{(1)}$) and plate bottom ($A^{(2)}$) reflections change with number of cycles. It is interesting to compare the signals [Figs. 6(c) and (d)] taken at the same number of cycles but only under different load levels. While the reflection signal recorded at 2.3-kN load changes significantly as the number of cycles increases, no change is observed until 83% of the fatigue life, when the signal is recorded at 0.3-kN load. The gates for the first and plate bottom reflection signals are shown.

Figure 7(a) shows peak-to-peak amplitudes $A_N^{(1)}$ of the first reflection normalized by the signal amplitude before the fatigue test $A_o^{(1)}$ as a function of the number of cycles N. This dependence is actually interpreted as ultrasonic reflected amplitude versus crack length, which grows in complicated fashion with number of fatigue cycles.[20–23] The ultrasonic signals at different load levels were normalized with the same quantities $A_o^{(1)}$, since they depend only negligibly on the load. The first reflection amplitude has minima at 85 000–115 000 cycles depending on the static load as shown in Fig. 7(a). Since the crack emanates approximately from the midplane of the pit, the path difference of waves reflected from the pit front and crack surface is about half the surface wave wavelength at 5 MHz.[21] Therefore, the interference of reflected waves from pit front and crack surfaces is destructive, resulting in decrease of the first reflection amplitude. The first noticeable change of the first reflection is observed at around 50 000 cycles as marked by arrow (A) in Fig. 7(a). However, it was found that the crack initiates much earlier.[20,23]

Figure 7(b) shows the plate bottom reflections $A_N^{(2)}$ versus number of cycles N normalized by the signal amplitude before the fatigue test $A_o^{(2)}$. The amplitude of the plate bottom reflection increases slightly with the number of cycles since it is not affected by wave interference. At around 60 000 cycles the second reflection starts to increase rapidly. This occurs when the crack reaches the pit depth, substan-

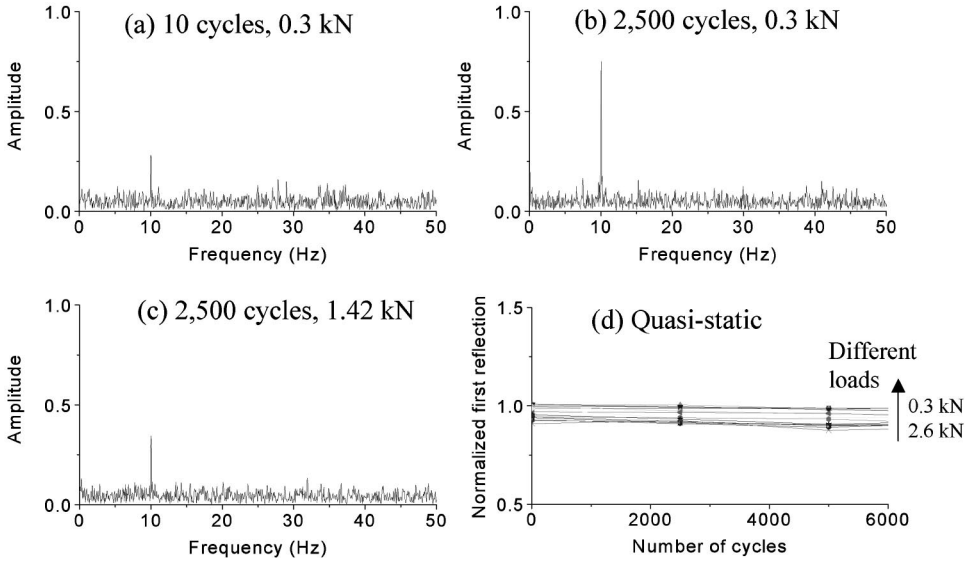


FIG. 9. Frequency spectra of modulation amplitudes (a), (b), and (c) at 10 and 2500 fatigue cycles for different loads. (d) is the normalized reflection amplitude from quasistatic measurements versus number of fatigue cycles in 0.3–2.6-kN load range.

tially increasing surface wave transformation to shear wave, thus increasing the plate bottom signal. Therefore, one can expect crack depth to be about pit depth at around 80 000 cycles. As observed in Figs. 6 and 7, the ultrasonic signals are sensitive to the load level at which they are measured. This phenomenon is related to the opening–closing behavior of the crack.[20,23]

In Fig. 8(a), the normalized first ($A_N^{(1)}$) and plate bottom ($A_N^{(2)}$) reflections versus cyclic load are shown at different numbers of cycles for the Al-2024-T3 sample. The first and bottom reflection signals were normalized with those signals before fatigue, $A_o^{(1)}$ and $A_o^{(2)}$, since they depend only negligibly on the load. As can be seen in Fig. 8(a), the reflected signal amplitude passes through a minimum with increase of load, which is due to the interference between the pit and the crack reflected signals. The signal interferences enhance sensitivity to determine crack initiation and crack mouth opening; however, this complicates the determination of the crack-fully opening loads.

The plate bottom reflections versus cyclic load at different fatigue cycles are shown in Fig. 8(b). The behavior has a common trend: each curve after a plateau begins to bend upward at a certain load level and saturates at some higher load. The bend-up load level occurs when the crack opening/closure boundary reaches the pit depth; it is slightly above the crack mouth opening (closure) load. The saturation of the reflection-versus-load curve occurs at a fully open crack. The behavior changes significantly at the end of fatigue life; for example, at 150 000 cycles, just before sample failure, the reflection amplitude does not change with varying load level since the crack tip deforms as a plastic hinge and the crack remains open during the whole loading cycle. The plate bottom reflection is useful in determining the crack-fully open load and, for high numbers of cycles, the load slightly above the mouth-opening load.

The transmitted signal has also been recorded in our experiments and it contains useful information; however, the results will be reported elsewhere in conjunction with other experiments.

B. Modulation measurements

1. Modulation results and discussion

As an example, Fig. 9 shows the frequency spectra of the modulated first reflection signals at 10 and 2500 cycles. Small background modulation is observed at ten cycles [Fig. 9(a)]. This is possibly due to slight deformation of the surface cavity shape under modulation load and corresponding change of the reflected signal. This modulation level is considered a background reference. Higher level of modulation is observed at 2500 cycles [Fig. 9(b)] at 0.3-kN static load. This change is due to an initiated crack which is smaller than 40 μm.[21] Since the crack initiated in the cavity-induced plastic-yielding zone, the compressive stress by the unrestored plastic deformation[20] tends to close the crack, as shown in Figs. 3 and 8. Therefore, at a low static load (0.3 kN) the crack is closed; thus, the crack opening–closure and the change of interfacial stiffness at contacting crack surfaces are responsible for the modulation. At a higher load (1.42 kN) the crack is fully open and the modulation amplitude decreases to a level close to the background modulation level [Fig. 9(c)]. For comparison, the normalized reflection amplitude versus number of cycles for different static loads is shown in Fig. 9(d). It shows no sensitivity to crack initiation at this number of cycles. The small load dependence is due to the slight deformation of the cavity.

Figure 10(a) shows the dependences of the modulation amplitudes of the first reflection signals versus the mean static load for several numbers of fatigue cycles. The number of fatigue cycles is related to the crack length.[21] The maximum modulation occurs at static loads in the middle range except for 2500 cycles. The crack at 2500 cycles is so small that it is easily opened and modulation reduced with the load. This decline is similar to the decline of the right (after maximum) part of the curves [Fig. 10(a)] for higher number of cycles [larger cracks]. This is important since higher modulation is obtained on a small closed crack with a smaller static load. As the closure (opening) stress changes with the crack growth,[20] the maximum modulation occurs at different static load levels: e.g., 1.7 kN at 50 000 cycles and 1.96 kN

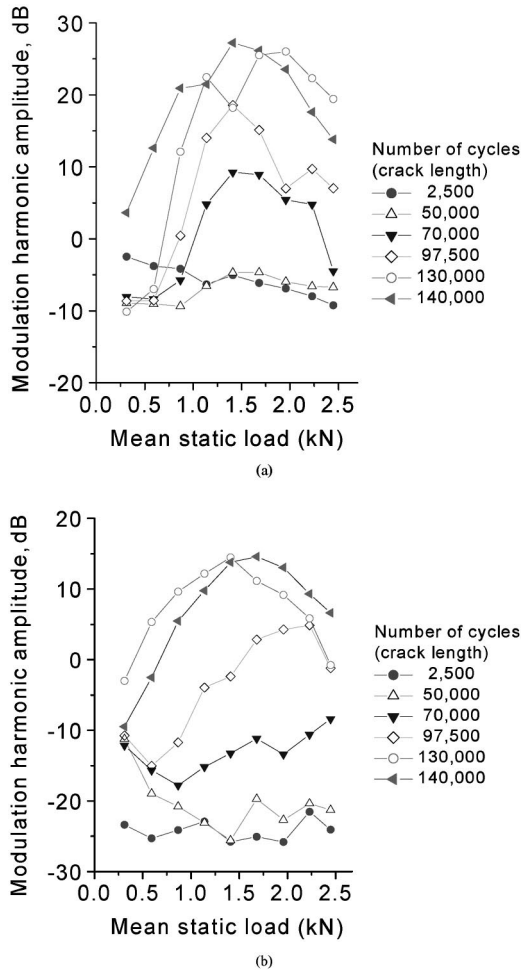


FIG. 10. Modulation amplitudes of the reflection signals versus the mean static load for different numbers of cycles (crack length). (a) the first reflection; (b) the plate bottom reflection.

at 130 000 cycles. As explained below, the modulation is maximized at static loads where the crack opening rate is highest with the load. It is noted that the modulation amplitude at lower static loads for the small crack developed at 2500 cycles is slightly higher than those for a larger grown crack at a higher number of cycles. The mechanism of this phenomenon is not completely clear; however, it may be due to the flattening effect of slightly unmated contacting asperities on the crack surfaces during repeated rubbing of the crack surfaces and according increase of the interfacial stiffness. Note that this change is much smaller than the change due to crack opening. At 140 000 cycles, the crack has almost completely transited through the sample relieving clamping compression stresses; for this reason it becomes open even at the lowest loads.

Figure 10(b) shows the modulation amplitudes of the plate bottom reflection signals versus mean static load for different numbers of cycles. As mentioned in the discussion of Fig. 7(b), the crack reaches the pit depth at around 60 000 cycles. Before this, the bottom reflection is formed by the mode-converted shear wave launched at the pit bottom and thus is not significantly affected by the crack growing at the top of the pit. Therefore, the modulation amplitude before this number of cycles is very small (<-10 dB). The difference in the modulation sensitivity of a bulk inhomogeneity and of a small crack is roughly -20 dB. Until 70 000 cycles, the level of modulation amplitude is under background noise (level -10 dB), which indicates that the bottom reflection signal is not appropriate for detecting small cracks by the modulation method.

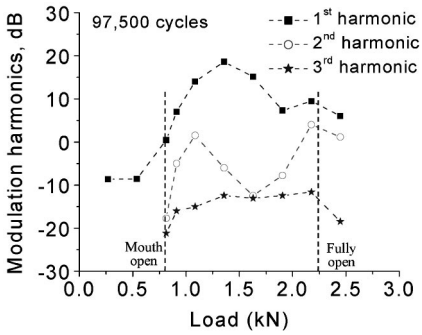


FIG. 11. Amplitude of modulation harmonics versus the mean static load at 97 500 cycles.

Figure 11 shows the dependence of first and higher modulation harmonics of the first reflection signal on the static crack opening load at 97 000 cycles. The crack opening loads are also shown. They are determined from the plate bottom reflection signals as indicated in Fig. 8(b) (see Ref. 20). The first harmonic (modulation amplitude) shows a rapid increase with crack mouth opening at around 0.75 kN and reaches its maximum at 1.3 kN. The modulation level increase (8 dB) at 0.75 kN demonstrates the sensitivity of the modulation method to the static load, compared with no measurable change of the reflected amplitude at the same loads (see Fig. 8). The increased modulation at loads higher than 0.75 kN is due to the open crack segment length change, which produces a stronger modulation effect than the interfacial condition change below 0.75 kN. The local minimum of the first modulation harmonic at about 1.75 kN corresponds to minimum of the reflection amplitude versus load,[21] which is formed by the interference of the signals reflected from the crack and the pit [see Fig. 8(a) for 90 kcycles].

The second harmonic peaks appear in the modulation spectrum at loads around 1.1 and 2.2 kN, which are slightly above the crack-mouth opening (0.8 kN) and slightly below the crack-fully opening (2.3 kN) loads. The aforementioned nonlinear crack opening/closure behavior at these loads is due to distortions of the modulated signal and is responsible for the considerable increase of the second harmonic amplitude. Further discussion of these experimental observations is given in the next section.

2. Discussion of modulation mechanism

The modulation harmonic amplitudes can be determined, for small modulation load amplitude, by series expansion of the surface wave reflection $R(F_o+f_\sim)$ with respect to its static equilibrium $R(F_o)$

$$\Delta R=(R(F_o+f_\sim)-R(F_o))$$
$$=R_o+R_1\cos\Omega t+R_2\cos 2\Omega t+\cdots, \quad (1)$$

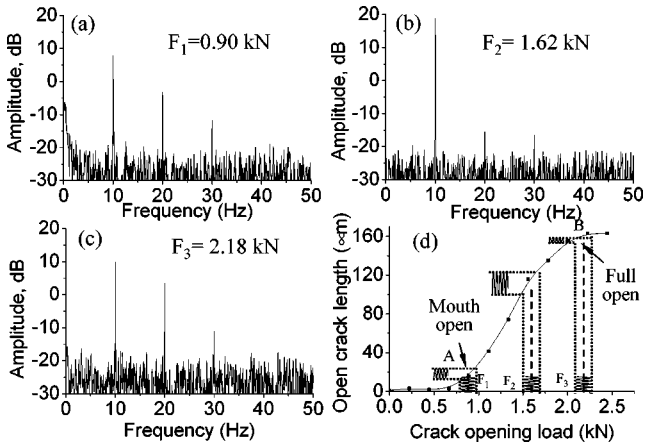


FIG. 12. (a), (b), (c) Frequency spectra of modulation amplitudes for different static loads at 97 500 cycles (crack size 330 μm), modulation frequency is 10 Hz. (d) Characteristic curve of open crack segment length versus static crack opening load. Points F_i indicate static loads at which low-frequency modulation load is applied. The corresponding modulation of the reflection coefficient is indicated schematically.

where $R_o = R_2 = R''(F_o)(\Delta f_{\sim})^2/4$ and $R_1 = R'(F_o)\Delta f_{\sim}$. The magnitude of the first modulation harmonic is proportional to the slope $[R'(F_o)]$ of the static reflection-versus-load characteristic curve (Fig. 8) and the modulation load amplitude Δf, and the second harmonic to the second derivative of the curve and (Δf).[2] When a crack is fully open, its size is not changed by the modulation load, and consequently the reflected or transmitted ultrasonic signal is weakly modulated. This occurs in the saturated region of the static reflection-versus-load characteristic curve (Fig. 8) and thus the slope $R'(F_o)$ is small. When the crack is tightly closed and nearly transparent to ultrasonic waves, again the slope $R'(F_o)$ is small and the modulation of the output signal is also small if the maximum tensile stress due to the modulation load is below the intrinsic compressive residual stress. The intermediate case shown in Fig. 3 corresponds to a crack state with part of its length open and part closed (this is a simplified view of a more realistic crack state having a transition region with a variable contact stiffness between the crack open and closed segments; the dependence of the contact stiffness on contact morphology and applied pressure is described elsewhere[8]). In this intermediate state, the length of the crack opening segment changes under modulation load and leads to related synchronous high modulation of the reflected ultrasonic signals.

To quantify this discussion, Figs. 12(a)–(c) show modulation spectra of the first reflection signal measured at different levels of static crack opening load for 330-μm crack depth (the crack size is estimated from the amplitude of the reflected ultrasonic signals in our similar experiments described in Ref. 20, and it grows to this size at 97 500 cycles). For a given crack length the amplitude of the first modulation harmonic depends strongly on the static load level and is maximum at the middle load [Fig. 12(b)]. A fresh fatigue crack has nearly perfectly conforming surfaces and thus the closed segment (l_c in Fig. 3) is nearly transparent to the ultrasonic probing waves. Therefore, the ultrasonic reflectivity is determined by the length of the open segment of the crack (l_o in Fig. 3) and the nature of the transition zone between the open and closed crack segments. The relative length of closed and open segments of the crack is determined by the balance of compressive residual and external tensile stresses which changes with applied modulation load, resulting in reflected ultrasonic signal modulation.

Figure 12(d) schematically illustrates the dependence of the open crack segment length on the static crack-opening load, which is determined from the reflection amplitude/static load characteristic curves (Fig. 8). The static load determines a modulation point on this nonlinear characteristic curve [modulation load and relative crack length modulation are indicated in Fig. 12(d) by small sinusoidal oscillations at loads F_1, F_2, F_3]. The maximum modulation is obtained at the maximum slope [Fig. 12(b)]. For static loads in the saturation regions [A and B in Fig. 12(d)] the first modulation harmonic is smaller [Figs. 12(a) and (c)].

The amplitude of higher modulation harmonics increases at static load close to the saturation regions (loads F_1 and F_3). As shown in Fig. 11 second harmonic peaks appear at loads 0.90 and 2.18 kN, which are slightly above the crack-mouth opening (0.8 kN) and slightly below the crack-fully opening (2.3 kN) loads. When the modulation load is applied at a static load (F_1) slightly above the crack-mouth-opening load, the crack during the modulation period undergoes transitions between the following states: tight closure↔loose closure on asperities at the crack mouth↔mouth opening↔small segment of crack opening. This nonlinear crack opening/closure behavior leads to distortions of the modulated signal and is responsible for the considerable increase of the second and third harmonic amplitudes. For modulation at a static load slightly below the crack-fully opening load [F_3 in Fig. 12(d)] the peak value of the sum of the static and modulation loads increases above the crack-fully opening load, leading to nonlinear crack opening behavior. This results in saturation of the ultrasonic reflectivity from the crack (reflectivity stops increasing with load) and increased levels of the higher modulation harmonics.

C. Dependence of modulation index on crack closure state

Figure 13 summarizes the crack-mouth-opening and crack-fully open loads versus number of fatigue cycles determined from each curve in Fig. 8(b). The crack-mouth-opening load is constant until around 112 500 cycles and then decreases, while the complete-opening load has its maximum at this number of cycles. At 112.5 kilocycles, the total crack length l is equal to the size l_p of the plastic-yielding zone. Crack-fully opening load is shown by open circles and crack-mouth-opening load by solid squares. H is the plate thickness; $l = H$ when the sample fails at 152.5 kilocycles. Different quantitative behavior of modulation is exhibited versus l/l_p: in regime I the crack is small and the crack-mouth- and crack-fully opening loads are not distinguishable; in regime II the crack is inside the plastic-yielding zone; and in regime III the crack tip departs from the plastic-yielding zone.

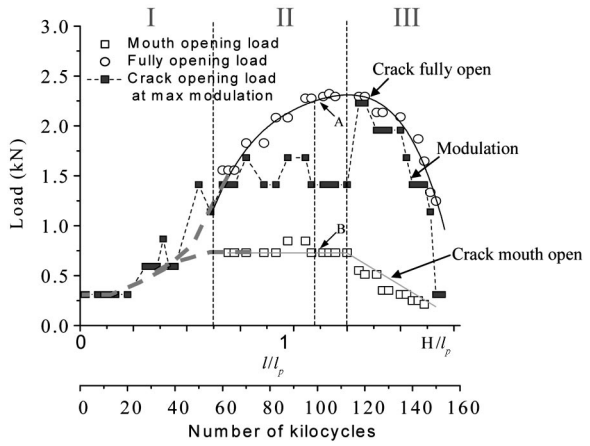


FIG. 13. Crack-mouth and fully opening loads versus number of fatigue cycles. Crack opening load that produces maximum modulation response (solid squares) versus relative total crack length (number of cycles) is also shown. At 112.5 kilocycles, the total crack length l is equal to the size l_p of the plastic-yielding zone. Crack-fully opening load is shown by open circles and crack-mouth-opening load by solid squares. H is the plate thickness: $l=H$, when the sample fails at 152.5 kilocycles. Different quantitative behavior of modulation is exhibited versus l/l_p: in regime I the crack is small and the crack-mouth and crack-fully opening loads are not distinguishable; in regime II the crack is inside the plastic-yielding zone; and in regime III the crack tip departs from the plastic-yielding zone.

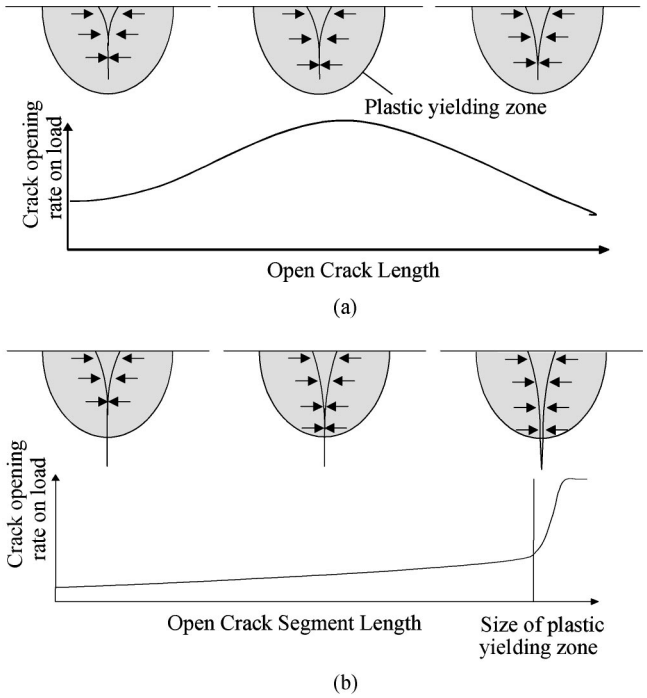


FIG. 14. Schematics to explain the jump of maximum modulation load in Fig. 10; they show the rate of the increase of open crack segment length versus its length. (a) The crack depth is smaller than that of the plastic-yielding zone. The whole length of the crack is under compressive stress and it is closed at low external tension load. The maximum crack-opening rate on load occurs when half the crack is open. (b) Crack is deeper that the plastic-yielding zone. The crack segment, which extends out of the plastic zone, is not constrained by compressive load. Thus, when the crack opening load is such that the open crack segment increases above the size of the plastic yielding-zone (indicated by a vertical line), the crack opens abruptly to its full size with a very small load increment and it has maximum opening rate producing maximum modulation amplitude.

The level of static load at which the first modulation harmonic reaches its maximum versus crack length is shown in Fig. 13. It is shown together with crack-mouth- and crack-fully opening loads determined from the measured ultrasonic amplitudes. Based on the analysis in Ref. 23, we have determined that the crack length reaches the cavity bottom (250 μm) at around 57 000 cycles and the plastic-yielding zone boundary (390 μm; see Fig. 3) at 112 000 cycles. At less than 57 000 cycles, the small crack is open under low static loads and the crack-mouth-opening and crack-fully opening loads are not distinguishable. When the crack length is comparable to the surface cavity size but still smaller than the plastic-yielding zone, the mouth-opening- and fully opening crack loads depart from one another. In this stage, while the crack is growing the maximum modulation occurs at loads near the crack-opening loads (following the dashed line) as shown in Fig. 13. When the crack becomes larger than the cavity depth, but is still in the plastic-yielding zone; the maximum modulation load becomes almost constant at load about 1.4 kN. In this stage the crack is constrained by nearly uniform compressive stress and the maximum modulation occurs when the crack is about half open ($l_o \sim l_c$; see Fig. 3), as shown schematically in Fig. 14(a). When the crack tip transits through the plastic-yielding zone border ($l_o + l_c = l_p$) one observes a jump of the static load value at which maximum modulation occurs. It increases to a level slightly below the crack-fully opening load. With crack growth out of the plastic-yielding zone the crack-fully opening load decreases due to reduction of the sample cross section and the corresponding reduction of compressive clamping stresses. The static load of maximum modulation follows the decrease of the crack-fully opening load. As the open crack segment grows out of the plastic-yielding zone, abrupt crack opening to full size occurs, since the crack front outside the plastic-yielding zone is not under compressive closure stress, leading to near equality of the crack opening and maximum modulation static loads (Fig. 13).

When the fully open crack is deeper than the plastic-yielding zone, the distribution of compressive stress acting on the crack surface changes as the crack is opened by the applied load. Until the open crack segment length reaches the plastic-yielding boundary, the whole part of the crack is constrained by the uniform compressive stress; thus, it becomes fully open gradually. As the open crack segment increases out of the plastic-yielding zone, the crack is opened suddenly because the remaining part of the crack is not under compressive load. For this reason, the maximum crack opening rate versus load always occurs after the open crack segment departs from the plastic yielding zone [Fig. 14(b)]. This occurs at loads just below the crack-fully opening loads (Fig. 13). The maximum change of the open crack segment length under the modulation load, i.e., the maximum modulation, occurs at maximum crack opening rate versus load.

IV. MODEL FOR SAW MODULATION FROM CRACK

We have simulated the modulation of the surface wave reflection signals due to modulation induced variations (Δl_o) of the open crack segment length (l_o in Fig. 3) using the low-frequency scattering model described in Ref. 21. It is

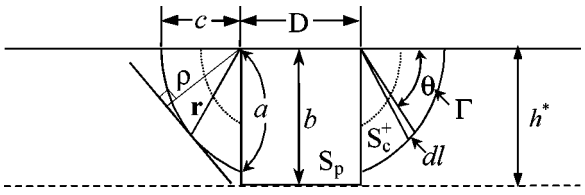


FIG. 15. Geometry of the cylindrical cavity with corner cracks (see also Fig. 3).

assumed that the closed part of the crack is transparent to ultrasonic waves and therefore the crack reflection is calculated for the open segment of the crack. The length of the crack open segment (l_o) is calculated as a function of opening static load (F_o) from the measured reflected amplitude without modulation (Fig. 8). Next, using this dependence, the modulation amplitude of the open crack segment length (Δl_o) is determined from the modulation load ($F_\sim$). Finally, the time-domain signal and spectrum of the modulated reflection amplitude are predicted using the scattering model.

The time-domain reflection signature from the cavity with the emanating crack is represented[21] by the sum of the signals reflected from the crack ($r_c(t,a)$) and the cavity ($r_p(t)$)

$$r(t,a)=r_p(t)+r_c(t,a) \tag{2a}$$

$$=r(t,a=0)+\int_{-\infty}^{\infty}R^c(\omega,a)V(\omega)e^{i\omega(t-D/V_R)}d\omega, \tag{2b}$$

where $R^c(\omega,a)$ is the normalized frequency-dependent reflection output of the receiver due to reflection from the crack with depth a, $V(\omega)$ is the frequency response of the measuring system, i is the imaginary unit, ω is the angular frequency, D is the cavity diameter, and V_R is the surface wave velocity. The cavity reflection coefficient in the presence of the crack is approximated by the cavity reflection coefficient without the crack $r(t,a=0)$, which can be measured prior to the fatigue.

To obtain the reflection amplitude from a crack with a pit, the reciprocity theorem[26,27] for elastic wave scattering is applied. Then, accounting for only the normal stress $\sigma_{zz}^{(II)}$ of the incident wave, which is dominant at shallow depths, the normalized reflection output signal from the crack is approximated as

$$R^c(\omega)=\frac{i\omega}{4P}\int_{S_c^+}\Delta u_z^{(I)}\sigma_{zz}^{(II)}dS, \tag{3}$$

where $\Delta u_z^{(I)}$ is the scattered displacement jump across the front (S_c^+) and back (S_c^-) crack surfaces; P is the input power applied to the transducer in the transmitting mode of operation. Equation (3) utilizes the symmetric configuration of two corner cracks at both sides of the pit. Using the results of Budiansky and O'Connell,[28] the crack reflection coefficient can be represented as a contour integral along the crack tip (Γ)

$$R^c(\omega)=\frac{i\omega(1-\nu^2)}{3EP}\int_\Gamma\rho(\mathbf{r})K_I^2\,dl, \tag{4}$$

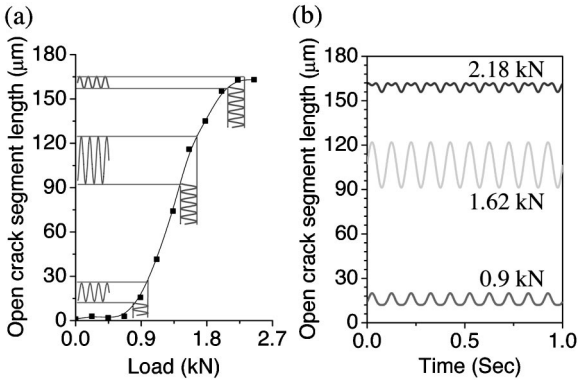


FIG. 16. Change of open crack segment length (OCSL) for modulation at different static load levels. Results obtained from ultrasonic reflection data at 87 000 fatigue cycles. (a) OCSL versus static load; (b) Change of OCSL with time by the modulation load at different static load levels.

where ν is Poisson's ratio, E is Young's modulus, K_I is the mode-I stress intensity factor of the corner crack, $\rho(\mathbf{r})$ is the perpendicular distance from the coordinate origin to the tangential line to the crack tip Γ at the given point $\mathbf{r}$, and dl is the line element on Γ as shown in Fig. 15.

To evaluate the stress intensity factor for this complicated geometry (Fig. 15), we approximate the stress ($\sigma_{zz}^{(II)}$) of the incident surface wave as a bending stress[21] in a plate with effective thickness h^*. Considering that the SAW stress[29] σ_{zz} changes its sign for aluminum at the depth $x/\lambda_R\approx0.225$, we take $h^*=0.45\lambda_R$ and the bending stress dependence as $\sigma_{zz}^{(II)}=\sigma_o(1-2x/h^*)$, where σ_o is the stress at the surface of the plate. In our experiment the ratio of pit depth b to the effective thickness h^* is $b/h^*\approx1$. So, we consider the h^* thick plate with a through-thickness hole as shown in Fig. 15. The mode-I stress intensity factor for symmetric corner cracks at a through-thickness hole in a plate subject to bending has been calculated numerically by Raju and Newman[30]

$$K_I=\sigma_o\sqrt{\frac{\pi a}{Q}}F_b\left(\frac{a}{h^*},\frac{a}{c},\frac{D}{h^*},\theta\right), \tag{5}$$

where a is crack depth, c is surface crack length, D is hole diameter, and Q is a function of the shape factor (a/c) of the crack. The boundary correction factor F_b is a function of the crack-and-hole geometries and the angle θ shown in Fig. 15. We use the numerical data of Raju and Newman[30] to calculate the stress intensity factors of different crack configurations. The aspect ratio (a/c, where c is crack length on the surface) is accounted for in the integration contour.

Finally, the normalized reflection signal from the crack is expressed as the line integral of the boundary correction factor along the front of the two corner cracks

$$R^c(\omega)=\frac{i\omega(1-\nu^2)\pi a\,\eta_T}{3EQ}\int_C\rho(\mathbf{r})F_b^2\left(\frac{a}{h^*},\theta\right)dl, \tag{6}$$

where $\eta_T=\sigma_o^2/P$ is the factor which is related to the electromechanical efficiency of the transducer.[21] The line integral is calculated numerically as a function of crack size with the known boundary correction factor and geometric parameters. Using Eqs. (2)–(6), the reflection coefficients versus the

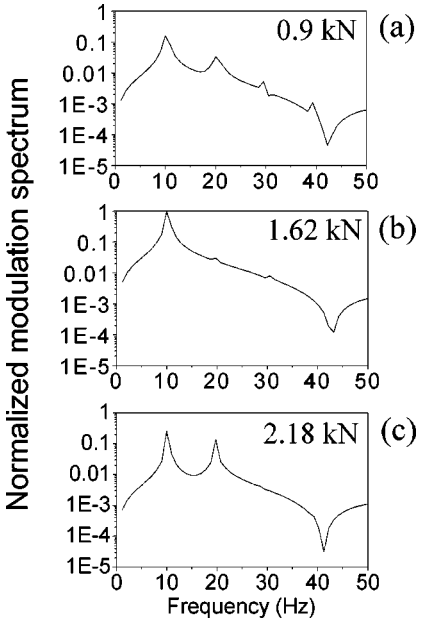


FIG. 17. Normalized modulation spectrum at 0.9, 1.62, and 2.18-kN static loads.

open crack segment length (OCSL) are obtained.

How changes of the modulation load transform to the oscillating open crack length segment is shown in Fig. 16(a) for three different static loads. OCSLs at static loads 0.9 and 2.18 kN show nonlinear distortion, while the OCSL at 1.62 kN shows nearly harmonic behavior. The modulation signal spectra at the three loads are shown in Fig. 17. At the static loads 0.9 and 2.18 kN, higher modulation harmonics appear since modulation occurs on the strongly nonlinear part of the characteristic curve. The second harmonic generation at these loads is supported by the experimental results shown in Fig. 14. At static load 1.62 kN, the OCSL characteristic versus load is almost linear and therefore the harmonic generation is not strong and the first harmonic has the largest amplitude because of the highest slope on the characteristic curve shown in Fig. 16(a).

Figure 18 shows comparison of the predicted and measured first and second modulation harmonics. The relative magnitudes of the modulation spectra are in quite close agreement, which is consistent with the proposed model for surface wave modulation by a partially closed crack.

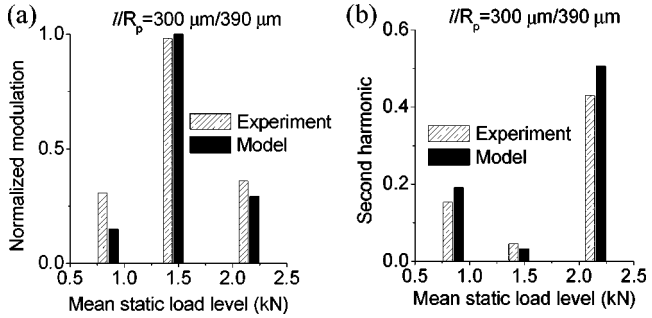


FIG. 18. Comparison of calculated and measured normalized modulation harmonic amplitudes for different mean static loads at 87 000 cycles. (a) First modulation harmonic $A^{(1)}$ normalized by the calculated value at 1.5 kN; (b) Second harmonic $A^{(2)}$ normalized by the $A^{(1)}$ at the same load.

V. CONCLUSION

It is shown that the modulation amplitude depends strongly on the crack length and the crack closure condition. For a closed fatigue crack under compressive stress, the modulation loading at a low mean static load can enhance considerably the modulation amplitude. The modulation is relatively weak when a crack is completely closed or completely open. There also exists an optimum static load at which the maximum modulation occurs. It corresponds to the load at which the crack opening rate versus load is a maximum. The change in optimum static load occurs abruptly when the fatigue crack grows larger than the plastic-yielding zone. Strong second harmonic generation is observed when the modulation is applied at loads near the crack-fully opening or the crack-mouth-opening loads. The response of the cracks under modulation and the generation of higher harmonics are predicted by the scattering model. We attribute the dominant mechanism of surface wave modulation by a partially closed fatigue crack in the plastic-yielding zone to the complex phenomenon of the open crack segment length change with modulation load. In this model we assumed abrupt transition from open to closed crack segments; however, this transition is gradual and the underlying physics of this transition needs to be further investigated. The results demonstrate that the effectiveness of the modulation technique may vary due to the dominating influence of the crack closure state; however, detectability of small closed cracks is significantly enhanced. The model and experimental results are in good agreement.

ACKNOWLEDGMENT

This work was partially sponsored by the Federal Aviation Administration (FAA) under Contract No. 97-C-001.

[1] O. Buck, W. L. Morris, and J. M. Richardson, "Acoustic harmonic generation at unbonded interfaces and fatigue cracks," Appl. Phys. Lett. **33**, 371–373 (1978).
[2] I. Y. Solodov, "Ultrasonics of nonlinear contacts: Propagation, reflection, and NDE-applications," Ultrasonics **36**, 383–390 (1998).
[3] S. Hirose and J. D. Achenbach, "Higher harmonics in the far field due to dynamic crack-face contacting," J. Acoust. Soc. Am. **93**, 142–147 (1993).
[4] L. A. Ostrovsky and P. A. Johnson, "Dynamic nonlinear elasticity in geomaterials," Riv. Nuovo Cimento **24**, 1 (2001).
[5] T. Meurer, J. Qu, and L. J. Jacobs, "Wave propagation in nonlinear and hysteretic media—A numerical study," Int. J. Solids Struct. **39**, 5585–5614 (2002).
[6] V. Gusev, A. Mandelis, and R. Bleiss, "Theory of strong photothermal nonlinearity from sub-surface non-stationary (breathing) cracks in solids," Appl. Phys. A: Mater. Sci. Process. **A57**, 229 (1993).
[7] I. Y. Solodov and B. A. Korshak, "Instability, chaos, and memory in acoustic-wave–crack interaction," Phys. Rev. Lett. **88**, 014303–014305 (2002).
[8] A. Baltazar, S. I. Rokhlin, and C. Pecorari, "On the relationship between ultrasonic and micromechanical properties of contacting rough surfaces," J. Mech. Phys. Solids **50**, 1397–1416 (2002).
[9] V. Zaitsev, V. Gusev, and B. Castagnede, "Thermoelastic mechanism for logarithmic slow dynamics and memory in elastic wave interactions with individual cracks," Phys. Rev. Lett. **90**, 075501-1–075501-4 (2003).
[10] J. A. TenCate, E. Smith, and R. A. Guyer, "Universal slow dynamics in granular solids," Phys. Rev. Lett. **85**, 1020–1023 (2000).
[11] A. Moussatov, V. Gusev, and B. Castagnede, "Self-induced hysteresis for nonlinear acoustic waves in cracked material," Phys. Rev. Lett. **91**, 124301-1–124301-4 (2003).
[12] H. Xiao and P. B. Nagy, "Enhanced ultrasonic detection of fatigue cracks

by laser-induced crack closure,'' J. Appl. Phys. **83**, 7453–7460 (1998).
[13] Z. Yan and P. B. Nagy, ''Thermo-optical modulation for improved ultrasonic fatigue crack detection in Ti-6Al-4V,'' NDT & E Int. **33**, 213–223 (2000).
[14] D. Donskoy, A. Sutin, and A. Ekimov, ''Nonlinear acoustic interaction on contact interfaces and its use for nondestructive testing,'' NDT & E Int. **34**, 231–238 (2001).
[15] K. E-A. Van Den Abeele, P. Johnson, and A. Sutin, ''Nonlinear elastic wave spectroscopy (NEWS) techniques to discern material damage, I. Nonlinear wave modulation spectroscopy (NWMS),'' Res. Nondestruct. Eval. **12**, 17–30 (2000).
[16] K. E-A. Van Den Abeele, A. Sutin, J. Carmeliet, and P. A. Johnson, ''Micro-damage diagnostics using nonlinear elastic wave spectroscopy,'' NDT & E Int. **34**, 239–248 (2001).
[17] V. V. Kazakov, A. Sutin, and P. A. Johnson, ''Sensitive imaging of an elastic nonlinear wave-scattering source in a solid,'' Appl. Phys. **81**, 646–648 (2002).
[18] K. J. Parker, S. R. Huang, R. A. Musulin, and R. M. Lerner, ''Tissue response to mechanical vibrations for sonoelasticity imaging,'' Ultrasound Med. Biol. **16**, 241–246 (1990).
[19] J. Ohpir, F. K. Garra, E. Konofagou, T. Krouskop, R. Righetti, and T. Varghese, ''Elastographic imaging,'' Ultrasound Med. Biol. **26-S1**, S23–S29 (2000).
[20] S. I. Rokhlin and J.-Y. Kim, ''*In situ* ultrasonic measurement of crack closure,'' Int. J. Fatigue **25**, 51–58 (2003).
[21] J.-Y. Kim and S. I. Rokhlin, ''Surface acoustic wave measurements of small fatigue cracks initiated from a surface cavity,'' Int. J. Solids Struct. **39**, 1487–1504 (2002).
[22] J.-Y. Kim, V. A. Yakovlev, and S. I. Rokhlin, ''Parametric modulation mechanism of surface acoustic wave on a partially closed crack,'' Appl. Phys. Lett. **82**, 3202–3204 (2003).
[23] S. I. Rokhlin and J.-Y. Kim, ''*In situ* ultrasonic monitoring of surface fatigue crack initiation and growth from surface cavity,'' Int. J. Fatigue **25**, 41–49 (2003).
[24] S. I. Rokhlin and J.-Y. Kim, ''Surface acoustic wave characterization of pitting corrosion damage with fatigue cracks,'' in *Nondestructive Materials Characterization with Application to Aerospace Materials*, edited by N. G. H. Meyendorf, P. B. Nagy, and S. I. Rokhlin (Springer, New York, 2004), pp. 142–179.
[25] S. I. Rokhlin, J.-Y. Kim, H. Nagy, and B. Zoofan, ''Effect of pitting corrosion on the fatigue crack initiation and fatigue life,'' Eng. Fract. Mech. **62**, 425–444 (1999).
[26] G. S. Kino, ''The application of reciprocity theory to scattering of acoustic waves by flaws,'' J. Appl. Phys. **49**, 3190–3199 (1978).
[27] B. A. Auld, ''General electromechanical reciprocity theory relations applied to the calculation of elastic wave scattering coefficients,'' Wave Motion **1**, 3–10 (1979).
[28] B. Budiansky and R. J. O'Connell, ''Elastic moduli of a cracked solid,'' Int. J. Solids Struct. **12**, 81–97 (1976).
[29] I. A. Viktorov, *Rayleigh and Lamb Waves—Physical Theory and Application* (Plenum, New York, 1967).
[30] I. S. Raju and J. C. Newman, Jr., ''Stress-intensity Factors for Two Symmetric Corner Cracks,'' in *Fracture Mechanics*—ASTM STP 677, edited by C. W. Smith (ASTM, Philadelphia, 1979), pp. 411–430.

Numerical and experimental analysis of second-order effects and loss mechanisms in axisymmetrical cavities

Cleofé Campos-Pozuelo[a] and Luis Elvira-Segura
Instituto de Acústica, CSIC, Serrano, 144, 28006 Madrid, Spain

Bertrand Dubus[b]
Institut d'Electronique de Microélectronique et de Nanotechnologie, Département ISEN, UMR CNRS 8520, 41 boulevard Vauban, 59046 Lille Cedex, France

(Received 20 February 2003; revised 10 December 2003; accepted 29 December 2003)

This paper deals with the analysis of finite amplitude acoustic waves in three-dimensional resonant cavities. A specific finite element model is proposed which includes: (i) the pressure field nonlinearity using a perturbative method; (ii) the loss mechanisms using an experimentally determined effective bulk attenuation or modeling viscous and thermal losses at the walls and acoustic radiation through the aperture with complex impedances. An axisymmetrical cavity with transversal dimension larger than the wavelength has been experimentally studied. The acoustic field is generated by a high-power flexurally vibrating transducer which generates high pressures. Measurements are performed for the fundamental and second-order pressure components for several cavity resonance modes. An important standard deviation is observed. Experimental data are compared to predicted pressure field distributions. Numerical models describe the pressure distribution correctly but overestimate the amplitude. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1687734]

PACS numbers: 43.25.Gf [MFH] Pages: 1973–1981

I. INTRODUCTION

The framework of this paper is the applications of ultrasonic energy in industrial processing. A number of industrial processes can be improved by using high power ultrasound (agglomeration,[1] particle filtration,[2] food processing,[3] etc.). The acoustic field inside the processing chamber and the efficiency of the system transducer-chamber are determinant factors for the feasibility of the applications. At present numerical models are widely used to model three-dimensional linear reactors, including fluid-structure coupling.[4] Nevertheless the development of practical systems requires a better knowledge of the nonlinear pressure field inside the reactor. A number of papers have been published on nonlinear standing waves. Saenger and Hudson[5] developed some experiments in one-dimensional rigid tubes and an analytical model explaining their results. The classical theoretical work on this subject is the analytical approach of Chester[6] referring to one-dimensional rigid walls cavities. In this approach[6] bulk attenuation is considered and a narrow tube approximation[7] is used to model losses due to the boundary layer. In the last several years several papers have been published dealing with numerical models solving the one-dimensional equations for a rigid tube.[8–10] These models work with strongly nonlinear plane waves and only bulk attenuation is considered. In Refs. 8 and 10 changes of section are taken into account but always in a one-dimensional approach. Nevertheless literature referring to experimental characterization and modeling of finite amplitude standing waves in real cavities is scarce. On the other hand, in any real situation thermal and viscous losses in the cavity walls must be considered in addition to the mainstream losses. Coppens and Sanders[11] presented an analytical model for three-dimensional cavities. Their model requires empirical values of the cavity quality factors at the different frequencies. They also observed that the agreement between experimental results and theoretical predictions decays when there are degeneracies. The purpose of this work is to improve the Finite Element Method (FEM) algorithm developed previously[12] to describe finite amplitude sound fields in three-dimensional close domains where attenuation effects in the boundaries are relevant. Two physical effects are more specifically analyzed.

(1) The acoustic field nonlinearity: A finite element modeling, previously developed for the description of progressive finite amplitude waves,[12] is used. It combines a perturbative method to linearize the equations and a classical finite element method. Second-order approximation is the only restriction of the method.
(2) The loss mechanisms in the resonant cavity: Specific local impedance finite elements are introduced, in addition to usual bulk viscous losses, to describe the viscous and thermal losses at the walls and the acoustic radiation from the cavity aperture. An alternative technique using only an effective bulk attenuation determination is also considered.

Measurements of the fundamental and second-harmonic field inside an axisymmetrical chamber were done to compare numerical predictions with nonlinear real acoustic fields. According to the authors, it is the first work where nonlinear acoustic measurements of the Sound Pressure

[a]Electronic mail: ccampos@ia.cetef.csic.es
[b]Electronic mail: bertrand.dubus@isen.fr

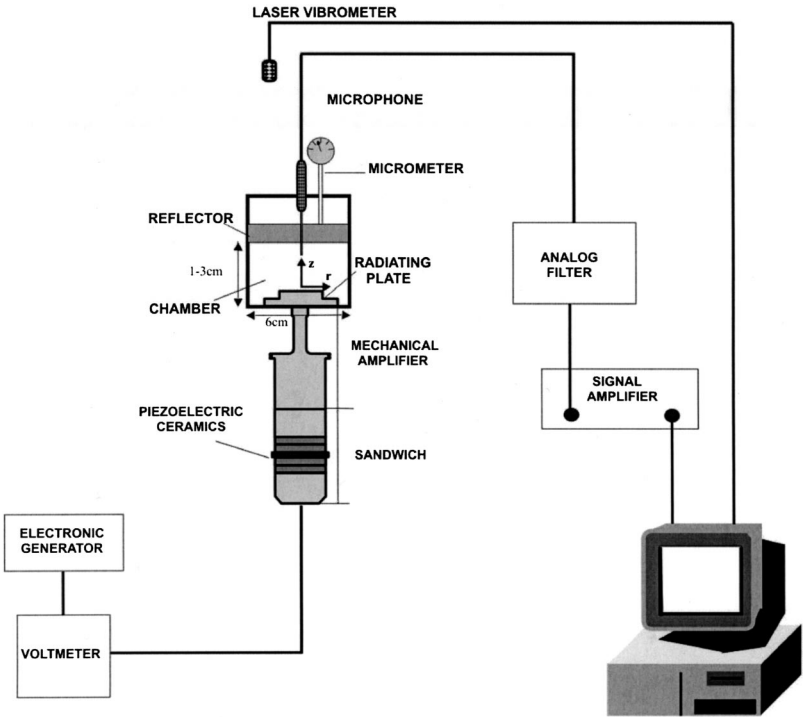


FIG. 1. Flexural stepped plate transducer, resonant chamber, and experimental setup.

Level (SPL) distribution inside an axisymmetrical cavity (not only at boundaries) are presented.

II. THEORY

A. Fundamental equations

The study considers standing waves of finite but moderate amplitude in a homogeneous fluid. Only second-order terms in the acoustic Mach number are considered. Lagrangian coordinates are used. In order to obtain a three-dimensional second-order wave equation, the following Taylor series expansion of the isentropic equation of state is considered:

$$\rho=\rho_0+a_1(p-p_0)+\frac{a_2}{2}(p-p_0)^2, \tag{1}$$

where p is the pressure, p_0 is the pressure at rest, ρ is the fluid density, ρ_0 is the fluid density at rest, $a_1=(\partial\rho/\partial p)_{S,p_0}$, $a_2=(\partial^2\rho/\partial p^2)_{S,p_0}$, S is the entropy. For an ideal gas $a_1=1/c_0^2$ and $a_2=-(\gamma-1)/\rho_0c_0^4$, c_0 is the small-signal sound speed and γ is the ratio of specific heats.

Combining state equation (1) with equations expressing the conservation of mass and momentum in Lagrangian coordinates, the following second-order wave equation is obtained:[12]

$$-\frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}+\nabla^2 p=-\frac{\beta}{\rho_0c_0^4}\frac{\partial^2(p-p_0)^2}{\partial t^2}, \tag{2}$$

where

$$\beta=\rho_0c_0^4\left(\frac{a_1^2}{\rho_0}-\frac{a_2}{2}\right)$$

is the nonlinearity parameter of the fluid (for an ideal gas $\beta=(\gamma+1)/2$) which determines the nonlinear characteristic of the propagation medium. For common fluids, the value of this parameter can be found in the literature (see Ref. 13, for example).

To solve the second-order equation (2) the successive approximations method is applied. This method assumes a solution for the acoustic pressure consisting of the addition of two terms,

$$p-p_0=p_1+p_2, \tag{3}$$

where p_1 represents the first-order approximation and p_2 the second-order correction, p_2 being much smaller than p_1. Substituting Eq. (3) into Eq. (2), taking into account that p_1 is the linear solution, and neglecting all the terms of the third or higher order, the following equation is obtained for the second-order correction, p_2:

$$-\frac{1}{c_0^2}\frac{\partial^2 p_2}{\partial t^2}+\nabla^2 p_2=-\frac{\beta}{\rho_0c_0^4}\frac{\partial^2 p_1^2}{\partial t^2}. \tag{4}$$

Equation (4) is a linear nonhomogeneous equation which can be solved by classical methods. For the simplest cases (one-dimensional standing waves) it can be solved analytically.[14] However, when acoustical or ultrasonic systems of more complex geometry are considered a numerical evaluation becomes necessary. In the following, a harmonic time dependence ($e^{j\omega t}$) is assumed for the first-order term.

B. Dissipative effects

Ultrasonic processing chambers are driven at a frequency corresponding to a cavity resonance in order to obtain high pressure amplitudes. Under these conditions, the efficiency of the ultrasonically assisted process is mainly determined by the pressure amplitude in the cavity. Therefore all loss mechanisms must be carefully modeled in order to predict pressure amplitudes in the chamber at resonance.

The chamber geometry is presented in Fig. 1. Three physical mechanisms determine the acoustic energy dissipation in the chamber: (i) the classical energy dissipation of acoustic waves associated with the gas viscosity and thermal conduction for constant frequency waves; (ii) the energy loss at the gas–wall interfaces associated with the thermal and viscous boundary layers; (iii) the radiation of acoustic energy through the aperture between the transducer head-mass and the wall. Two strategies are used to include energy loss in the models.

(1) Model A considers an effective complex sound velocity (for the whole volume of gas). An effective attenuation coefficient which includes all three loss mechanisms for each resonance of the chamber is experimentally determined. According to the axisymmetrical modal theory for a chamber with rigid walls, the complex wave number of the different modes can then be written,[15]

$$k_{nm}=k_{nm0}(1-\overline{\alpha}(1+j)) = \sqrt{\left(\frac{n\pi}{L}\right)^2+\left(\frac{\delta_{Om}\pi}{R}\right)^2} \times\left(1-(1+j)\frac{\sigma k_0\left(\frac{2}{L}+\frac{1}{R}\right)}{\left(\frac{n\pi}{L}\right)^2+\left(\frac{\delta_{Om}\pi}{R}\right)^2}\right), \tag{5}$$

where (m,n) is the resonance order, σ is the susceptance at the boundary surface, k_0 is the acoustic wave number in free space, and $\alpha=k_{nm0}\overline{\alpha}$ is the attenuation coefficient. The value of this effective attenuation coefficient which includes all three loss mechanisms for each resonance of the chamber is experimentally obtained. The experimental procedure for measuring these coefficients is detailed in Sec. III B.

(2) Model B includes distinct models for each mechanism and only relies upon physical constants from handbooks. The loss mechanism associated with the wave propagation in the volume is described by a complex sound velocity. The effect of the boundary layer on the system response is based on the classical theory of modal superposition.[16] The rigid wall condition is substituted by a local impedance condition (Z) to evaluate the effect of viscosity and thermal conduction. Equations for plane wave reflection at a solid surface are used for every solid boundary in the cavity:[16]

$$\frac{1}{Z}=\frac{1}{2}(1+j)\frac{\omega}{\rho_0 c_0^2}[l_{\mathrm{vor}}\sin^2\theta_i+(\gamma-1)l_{\mathrm{ent}}],$$

$$l_{\mathrm{vor}}=\sqrt{\frac{2\mu}{\omega\rho_0}},\quad l_{\mathrm{ent}}=\frac{l_{\mathrm{vor}}}{\sqrt{\mathrm{Pr}}}, \tag{6}$$

where μ is the fluid viscosity and Pr the Prandtl number. θ_i is the incidence angle for a plane wave reflection. The cavity resonance order is considered to determine an incidence angle on each cavity wall. In the case of a cylindrical cavity, the wave number $\mathbf{k}$ for the rigid cavity resonance of order (m,n) is written[15]

$$\mathbf{k}=k_r\mathbf{r}+k_z\mathbf{z}=\frac{\delta_{0m}\pi}{R}\mathbf{r}+\frac{n\pi}{L}\mathbf{z}, \tag{7}$$

where $\mathbf{r}$ and $\mathbf{z}$ are unitary vectors in the radial and the axial directions, R is the cavity radius, and L the cavity length. n and m are integers and δ_{0m} verify $J_0'(\pi\delta_{0m})=0$ where J_0 is the Bessel function of first kind and zero order. Thus, $\tan\theta_i=k_z/k_r=nR/\delta_{0m}L$ is used for surfaces perpendicular to the cavity radius and $\tan\theta_i=k_r/k_z=\delta_{Om}L/nR$ for surfaces perpendicular to the cavity axis. This approach relies upon an *a priori* evaluation of the incidence angle. It could be efficiently completed in the numerical model by an *a posteriori* correction of the incidence angle at each point of the wall surface as recently proposed by Bossart *et al.*[17]

The cavity aperture is modeled by using an acoustic radiation impedance condition. This acoustic impedance z at each point of the aperture is approximated by the average acoustic impedance of an annulus (external radius R_e, internal radius R_i) of uniform velocity mounted in a rigid baffle:[16]

$$z(\omega)=\frac{\rho_0 c_0}{\pi(R_e^2-R_i^2)} \times\left[\pi R_e^2\left(1-\frac{2J_1(2kR_e)}{2kR_e}+j\frac{2\mathbf{H}_1(2kR_e)}{2kR_e}\right) -\pi R_i^2\left(1-\frac{2J_1(2kR_i)}{2kR_i}+j\frac{2\mathbf{H}_1(2kR_i)}{2kR_i}\right)\right]. \tag{8}$$

J_1 (respectively, $\mathbf{H}_1$) is the the Bessel function of first kind (respectively, Struve function) of first order.

C. Numerical method

The finite element formulation of the fluid structure problem leads to the following set of equations for the linear term:[18]

$$\begin{bmatrix}[K]-\omega^2[M] & -[L]\\ -\rho_0^2c_0^2\omega^2[L]^T & [H]-\omega^2[M_1]\end{bmatrix}\begin{bmatrix}\mathbf{U}_1\\ \mathbf{P}_1\end{bmatrix}=\begin{bmatrix}\mathbf{F}\\ \mathbf{0}\end{bmatrix}, \tag{9}$$

where $[K]$ and $[M]$ are, respectively, the solid stiffness and consistent mass matrices, $\mathbf{U}_1$ is the nodal values vector of the solid displacement field (first-order term), $\mathbf{F}$ is the nodal values vector of the applied forces, $\mathbf{P}_1$ is the nodal values vector of the pressure field (first-order term), $[L]$ is the interface connectivity matrix, $[H]$ and $[M_1]$ are, respectively, the fluid stiffness and consistent mass matrices, and the superscript T means transposed.

To model the local impedances (aperture and walls), a mechanical impedance element is used. The acoustic impedance z at a point M on a surface S_p is defined as

$$p(M)-p_0=-p_e(M)=jz(\omega)\omega u_n(M). \tag{10}$$

u_n is the displacement normal to the surface and p_e the external pressure acting on the surface.[19] The nodal forces $\mathbf{R}^e$ due to the external medium and acting on surface S_p^e are written

$$\mathbf{R}^e = \int \int_{S_p^e} [N_p^e]^T p_e dS_p^e$$

$$= -j\omega z(\omega) \int \int_{S_p^e} [N_p^e]^T \mathbf{n}^T [N^e] \mathbf{U}^e dS_p^e$$

$$= -j\omega [Z^e(\omega)] \mathbf{U}^e, \tag{11}$$

where $\mathbf{n}$ is the unitary vector normal to the surface S_p^e, $[N_p^e]$ is the shape function matrix associated with the surface finite element, and $\lfloor Z^e \rfloor$ is the elementary mechanical impedance matrix. After assembling of $\lfloor Z^e \rfloor$ into $[Z]$ and combining with Eq. (8) the equation set for the linear term is written

$$\begin{bmatrix} [K] - \omega^2 [M] + j\omega [Z(\omega)] & -[L] \\ -\rho_0^2 c_0^2 \omega^2 [L]^T & [H] - \omega^2 [M_1] \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{P}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}. \tag{12}$$

For the second-order term, the finite element equations set of Ref. 12 is modified by including the mechanical impedance matrices and removing the matrices associated with the far-field radiation:

$$\begin{bmatrix} [K] - 4\omega^2 [M] + j2\omega [Z(\omega)] & -[L] \\ -4\rho_0^2 c_0^2 \omega^2 [L]^T & [H] - 4\omega^2 [M_1] \end{bmatrix} \begin{bmatrix} \mathbf{U}_2 \\ \mathbf{P}_2 \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{0} \\ \rho_0 c_0^2 \mathbf{R}_B \end{bmatrix}, \tag{13}$$

where $\mathbf{U}_2$ (respectively, $\mathbf{P}_2$) is the nodal values vector of the second-order pressure correction (respectively, displacement). $\mathbf{R}_B$ is a volumetric source term due to the fluid nonlinearity which is written

$$\mathbf{R}_B = \sum_e \mathbf{R}_B^e = -\sum_e \frac{2\omega^2 \beta}{\rho_0 c_0^4} \int \int \int_{\Omega_f^e} [N_p^e]^T \times [N_p^e] \mathbf{P}_l^e [N_p^e] \mathbf{P}_l^e d\Omega_f^e. \tag{14}$$

III. EXPERIMENTAL PROCEDURE

A. Experimental device

The experimental setup is displayed in Fig. 1.[20] A cylindrical air-filled chamber driven by a high power transducer is studied. The chamber was made of aluminum and has a diameter of 6 cm. The transducer consists of a circular plate of stepped shape driven at its center by a piezoelectric vibrator. The vibrator itself consists of piezoelectric rings (PZT4) in a sandwich arrangement and a solid horn, which acts as a vibration amplifier. The longitudinal vibration, generated by the transducer element and amplified by the horn, drives the radiating plate which vibrates flexurally in its first axisymmetrical mode. The transducer elements are calculated to be resonant at the working frequency (23 600 Hz). The vibrating system is driven at its resonance frequency by a power generator which incorporates a feedback circuit to automatically adjust the excitation frequency to the transducer resonance frequency.[21]

The length of the acoustic chamber is changed by moving the reflector by means of a precision screw (± 10 μm). This method is used to tune the cavity resonance to the transducer resonance. A small hole is made in the reflector for the velocity and pressure measurements. Laser interferometry (Polytec OFV-3000) is used to evaluate the vibrational velocity of the stepped plate. The acoustic pressure distribution and wave forms along the cavity axis are measured with a condenser microphone (Brüel and Kjær, 1/8 in.). The position of this microphone inside the chamber is PC-controlled. An analog frequency filter is used to quantify the harmonic content of the pressure at every measurement point.

B. Empirical evaluation of the attenuation coefficient

The quality factor of a resonance is usually determined by plotting the pressure variation versus frequency. This method is not suitable when the acoustic source is a narrowband transducer since the transducer band would perturb the measurement. We propose to determine the effective cavity attenuation by keeping the frequency constant and varying the tube length around the resonance length. We use a parameter that we call ''spatial quality factor,'' Q_s defined as the ratio between the tube length at resonance (maximum pressure), L, by the ''spatial bandwidth,'' $2\Delta L$, that is, the difference between the two values of the tube length at which the pressure is equal to the pressure at resonance divided by $\sqrt{2}$. The pressure variation (under the same hypotheses) is

$$\frac{p_L}{p_{L+\Delta L}} = \frac{\sqrt{\left(\frac{n\pi}{L}\right)^4 \left(\frac{2\Delta L}{L}\right)^2 + 4\alpha^2 \left(\left(\frac{n\pi}{L}\right)^2 + \left(\frac{\alpha_{0m}\pi}{R}\right)^2\right)}}{2\alpha \sqrt{\left(\frac{n\pi}{L}\right)^2 + \left(\frac{\alpha_{0m}\pi}{R}\right)^2}}. \tag{15}$$

From these results, the expression for the complex part of the wave number in the resonant chamber is obtained,

$$\alpha = \frac{1}{2Q_s} \left(\frac{n\pi}{L}\right)^2 \Big/ \sqrt{\left(\frac{n\pi}{L}\right)^2 + \left(\frac{\alpha_{0m}\pi}{R}\right)^2}, \tag{16}$$

where $Q_s = L/2\Delta L$. For a one-dimensional wave, the spatial quality factor $Q_s = n\pi/2\alpha L$ coincides with the typical quality factor obtained from the frequency plots.

C. Measurement procedure

The reflector wall is moved up to obtain a maximum pressure at the reflector and the modes are selected by comparing the measured length at resonance with a linear modal numerical analysis of the chamber.

First, measurements of the vibrational velocity of the transducer radiating plate are carried out by using a laser vibrometer for calibration and verification of the field excitation linearity. All measurements are made in working conditions, i.e., the laser beam reflected in the transducer center when the cavity is closed and tuned (see Fig. 1).

Second, the attenuation is evaluated from pressure measurements at the reflector for different tube lengths around the resonance length. For each mode, the measurement is performed for the fundamental and second-harmonic frequencies.

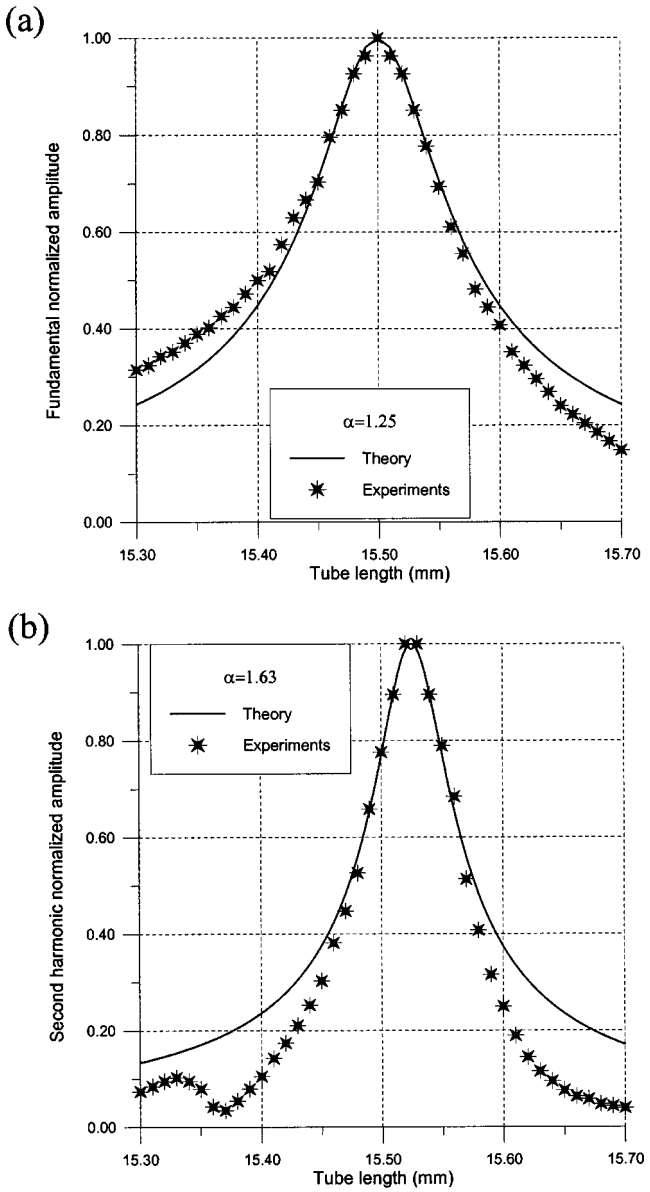


FIG. 2. Mode 1. (a) Fundamental amplitude at the reflector vs cavity length; (b) second-harmonic amplitude at the reflector vs cavity length. Dots: experiments; full line: fit of the theoretical model.

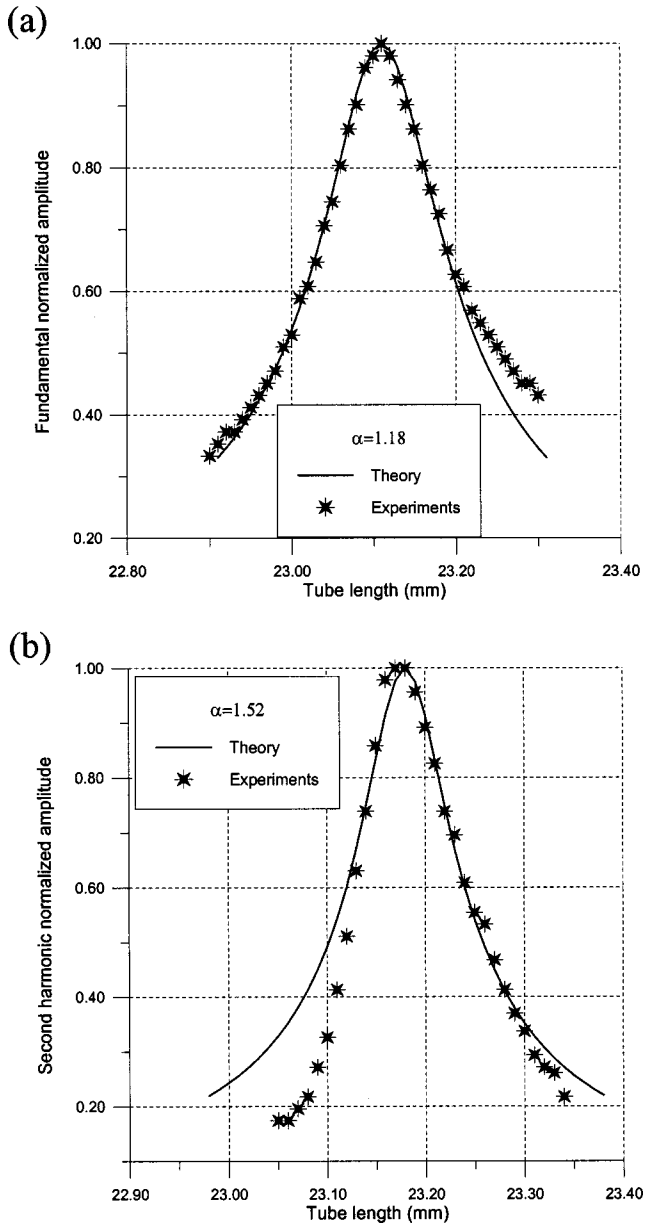


FIG. 3. Mode 2. (a) Fundamental amplitude at the reflector vs cavity length; (b) second-harmonic amplitude at the reflector vs cavity length. Dots: experiments; full line: fit of the theoretical model.

Finally, the fundamental and second-harmonic axial distributions of the acoustic field are evaluated by moving the microphone along the chamber axis for every tuned mode.

IV. RESULTS AND DISCUSSION

Experimental and numerical results for two cavity resonance modes are analyzed and compared in this section. The two measured modes, denoted, respectively, mode 1 and mode 2, correspond to resonance cavity lengths of 7.5 and 15.1 mm measured from the radiating plate center. If the cavity is approximated by an ideal cylinder, the modes correspond to $(n,m)=(2,1)$ and $(n,m)=(3,1)$, respectively.

The attenuation is evaluated following the procedure described in Secs. III B and III C for the fundamental and second harmonic of each mode. The resonance peaks are plotted in Figs. 2 and 3. The coefficients obtained from the resonance curves are displayed in Table I. Different attenuation coefficients are obtained for each mode whereas they are identical in an ideal cylinder [according to Eq. (16) with $\alpha_{0m}=0$]. The difference can be attributed to the real geometry of the chamber (not perfectly cylindrical, stepped profile of the transducer, pressure leakage). For each mode, the ratio of the second-harmonic attenuation coefficient to the fundamental attenuation coefficient is not very different from $\sqrt{2}$, value of this ratio in the case of a one-dimensional plane wave inside a tube with rigid walls. This result indicates that,

TABLE I. Effective attenuation coefficients measured for fundamental (α_f) and second-harmonic (α_h) frequencies.

Cavity length (mm)	α_f (m^{-1})	α_h (m^{-1})	α_f/α_h
7.5	1.25	1.63	1.304
15.1	1.18	1.52	1.288

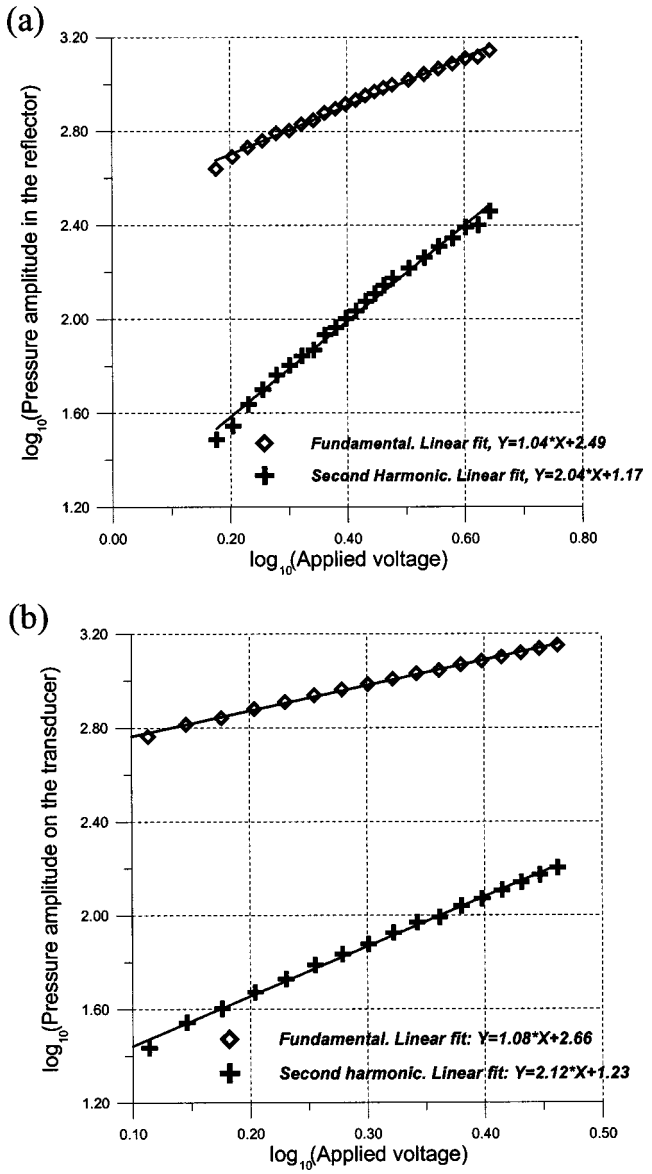


FIG. 4. Mode 1. Velocity amplitudes measured at the center of the transducer plate vs driving voltage. Diamonds: amplitude at the fundamental frequency; circles: amplitude at the second-harmonic frequency. Following measurements are made at a voltage of 2.25 V.

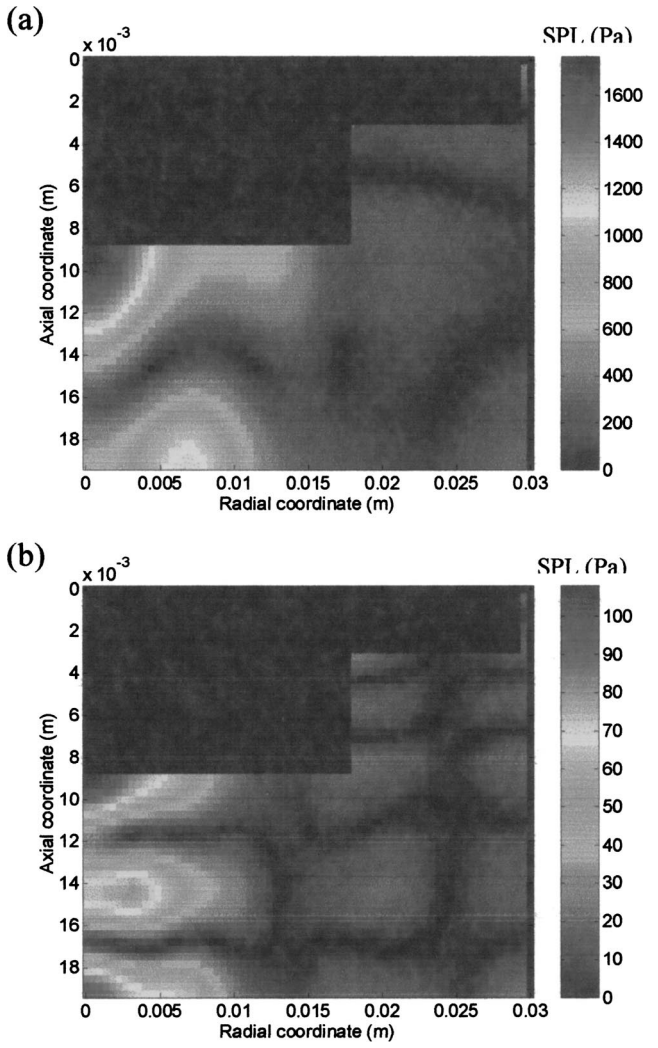


FIG. 5. Mode 1. Measured pressure amplitudes vs driving voltage. (a) Fundamental frequency; (b) second-harmonic frequency. Diamonds: at the transducer center; crosses: at the reflector center.

even in a wide tube ($r>2\lambda$), losses are mainly due to the boundary layers effect, for which a dispersion relationship of $\sqrt{\omega}$ is expected.

The particle velocity wave form at the center of the transducer has been analyzed in order to be sure of having harmonic excitation. Figures 4(a) and (b) show the experimental variation of the pressure fundamental and second-harmonic amplitude versus excitation at two different points on the cavity axis. A linear variation of the fundamental is observed while the second harmonic shows a parabolic increase with excitation. The linear behavior of the fundamental confirms that, for pressure amplitude ranges in which measurements are conducted, the energy transfer to harmonic frequencies is negligible. Therefore, the approximation assumed in Eq. (3) is valid (Fig. 4).

The finite element model is applied to analyze the different cavity modes. The transducer is represented by the radiating plate with a prescribed mechanical excitation at its center. The typical problem size is 4000 elements and 15 000 unknowns. The computation time is 30 s on PC (Pentium IV at 1.3 GHz, 256 MO RAM) to obtain solutions at fundamental and second-harmonic frequencies. Figures 5 and 6 show the calculated spatial distribution of the fundamental and second-harmonic components for each cavity length. The SPL distributions for both fundamental and second harmonic reveal the two-dimensional characteristic of the modes analyzed, where the axial and the radial components are present.

The experimental and numerical axial distributions are compared for the fundamental and second harmonic of both modes in Figs. 7 and 8. Transducer displacements of 0.52 and 0.85 μm are considered for the first and second mode, respectively. An important standard deviation (error bars) is observed, especially for SPL minima, where errors of 7 dB are reached. These deviations are lower near the SPL maxima (1–3 dB). These instabilities were also found by other authors measuring nonplanar modes in real cavities,[11,22] especially for the higher modes, where small changes in the propagating media, geometry, or boundaries cause important changes in the acoustic field. The standard

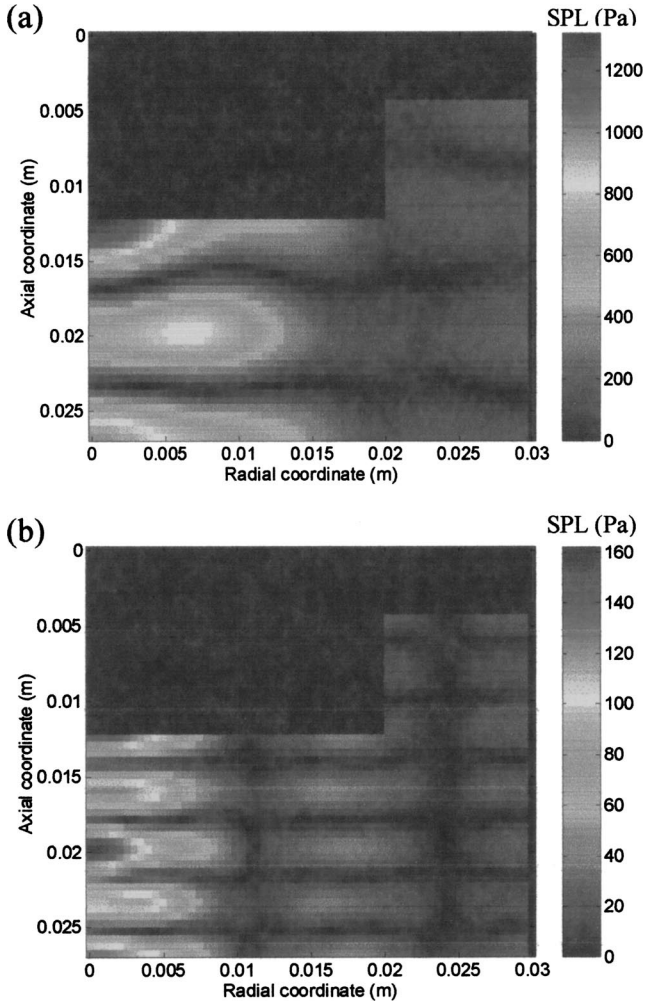


FIG. 6. Mode 1. Computed pressure amplitude isocontours. (a) Fundamental frequency; (b) second-harmonic frequency.

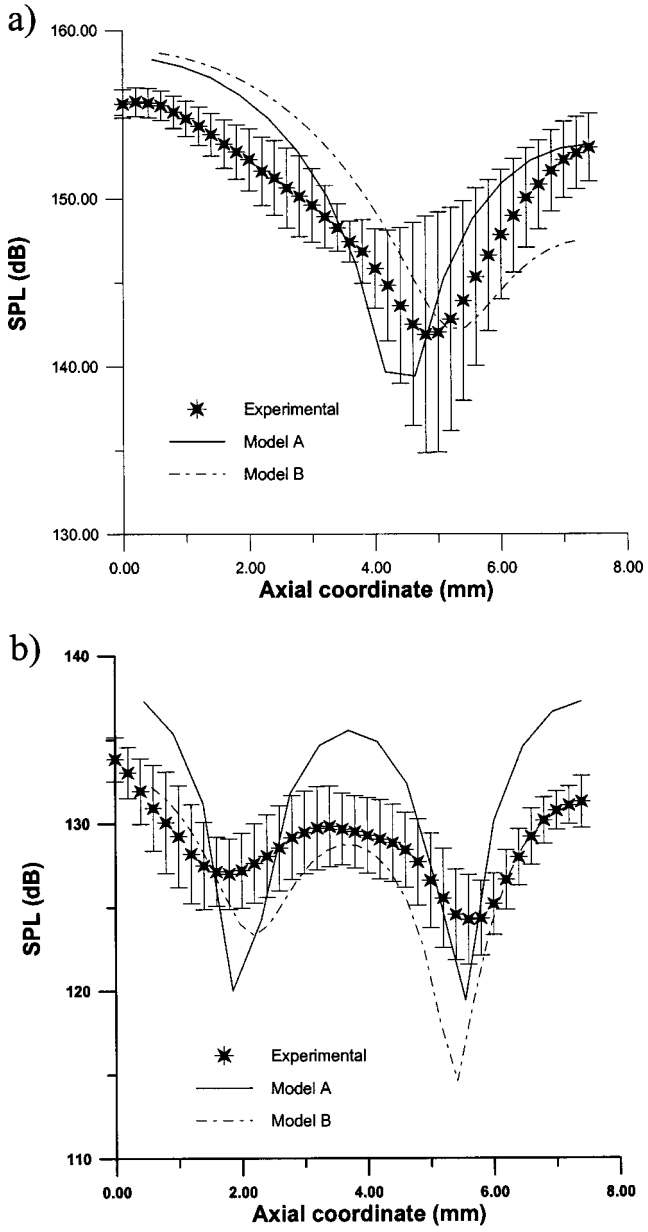


FIG. 7. Mode 1. Computed pressure amplitude isocontours. (a) Fundamental frequency; (b) second-harmonic frequency.

deviation increases for the second-harmonic component, as a result of its quadratic dependence with the fundamental and these different modes become closer as the frequency increases, leading to (near-) degeneracies.[11] This last effect can be observed in Fig. 2(b), where a small peak belonging to another mode appears for a 15,33 mm tube length.

Experimental data are compared with the two different numerical approaches based on models A and B. For the first mode (short cavity), the pressure distribution and amplitude are well predicted with a trend to smooth out the peaks and dips which can be attributed to an averaging effect of the microphone. For second mode (long cavity), both numerical models correctly describe the pressure distribution but overestimate the amplitude. Uncertainties on the aperture size are a possible explanation for this discrepancy. The only work found where absolute values of nonlinear measurements were obtained[23] gives errors up to 4 dB with the theory for a one-dimensional cavity at its end. For our three-dimensional cavity, the discrepancies between theory and experiments are of similar order, lying between 2 and 6 dB, except in the case of the second harmonic, model B and the second resonant mode (larger cavity), Fig. 8(b), where a 15 dB discrepancy was found. Probably, for that mode the attenuation effects in the chamber walls are more relevant and better described by model A where empirical evaluation of the coefficient is done. The agreement obtained between measurements and model A validates the attenuation measurement method and shows that losses at the boundary layer determine the dispersion relation for attenuation. Nevertheless the fairly good agreement obtained between measurement and model B shows that the pressure in resonant cavities of complex geometries can be predicted using only handbook values. These quantitative predictions are only possible if the complete geometry of the device is accurately known. These results also validate the second-order perturbation and its numerical formulation for close domains.

V. CONCLUSIONS

The pressure field at resonance in a three-dimensional cavity has been studied numerically and experimentally both at the fundamental frequency of the driving transducer and at

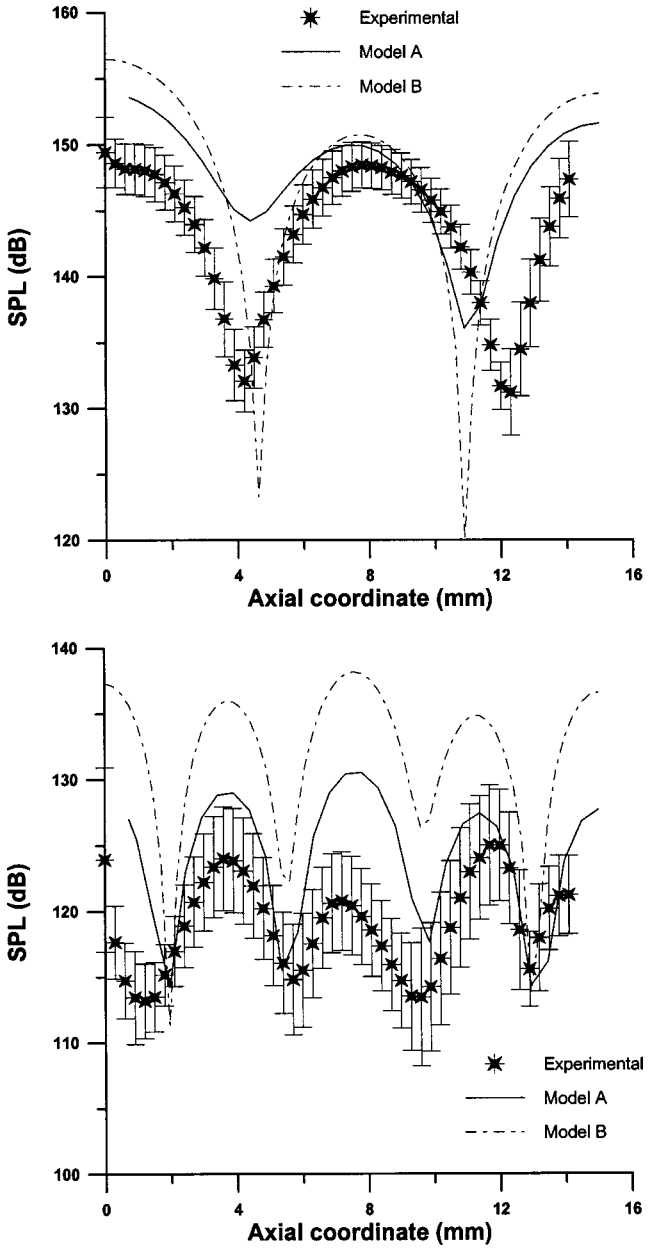


FIG. 8. (A) Pressure amplitude on the cavity axis. Mode 1 (a) Fundamental frequency, (b) second-harmonic frequency. Dots: measurement, full line: computation with model A; dashed line: computation with model B. Vertical bars show standard deviation on measured values. (B) Pressure amplitude on the cavity axis. Mode 2. (a) Fundamental frequency, (b) second-harmonic frequency. Dots: measurement, full line: computation with model A; dashed line: computation with model B. Vertical bars show standard deviation on measured values.

the second-harmonic frequency. In such resonant chambers, the attenuation is mainly determined by the viscous and thermal losses at the walls and the acoustic radiation from the aperture. Therefore, the modeling and identification of these losses has been the subject of specific attention. The main results obtained from this analysis are as follows.

(i) A method has been proposed to determine the effective attenuation coefficient of a cavity resonance with a narrow-band transducer. It is based on the pressure amplitude variation versus cavity length rather than frequency. When using an effective bulk attenuation coefficient, numerical predicted pressures in the cavity agree quantitatively with measured pressures, both for fundamental and second-harmonic frequencies.

(ii) Quantitative values of pressure amplitudes in the cavity at resonance can be predicted by using only handbook values to describe cavity losses (volumetric viscous losses, viscous and thermal losses at walls, acoustic radiation from aperture). Good agreement is found with measurements at fundamental and second-harmonic frequencies for two different cavity lengths.

(iii) The finite element formulation used to predict second-harmonic pressures was initially developed to model finite amplitude progressive waves.[12] The results presented in this paper validate this formulation for finite amplitude standing waves. More generally, this numerical modeling constitutes a powerful tool to analyze and design high power ultrasonic systems using resonant chambers.

ACKNOWLEDGMENTS

This work was supported by CNRS and CSIC (French-Spanish Cooperation Project No. 2002FR0004). The authors want to thank E. Andrés for technical support.

[1] E. Riera-Franco de Sarabia, J. A. Gallego-Juárez, V. M. Acosta-Aparicio, J. J. Rodríguez-Maroto, J. L. Dorronsoro, D. Sanz-Rivera, F. J. Gómez-Moreno, and M. Martín-Espigares, "Acoustic agglomeration of submicron particles in diesel exhausts: First results of the influence of humidity at two acoustic frequencies," J. Aerosol Sci. **31**, 827–828 (2001).
[2] J. J. Hawkes, W. T. Coakley, M. Gröschl, E. Benes, S. Armstrong, P. Tasker, and H. Nowotny, "Single half-wavelength ultrasonic particle filter: Predictions of the transfer matrix multilayer resonator model and experimental filtration results," J. Acoust. Soc. Am. **111**, 1259–1266 (2002).
[3] D. J. McClements, "Advances in the application of ultrasounds in food analysis and processing," Trends Food Sci. Technol. **6**, 93–99 (1995).
[4] J.-P. Morand and R. Ohayon, *Fluid-Structure Interaction: Applied Numerical Methods* (Wiley, New York, 1995).
[5] R. A. Saenger and G. E. Hudson, "Periodic shock waves in resonating gas columns," J. Acoust. Soc. Am. **32**, 961–970 (1960).
[6] W. Chester, "Resonant oscillations in closed tubes," J. Fluid Mech. **18**, 44–64 (1964).
[7] D. E. Weston, "The theory of the propagation of plane sound waves in tubes," Proc. Phys. Soc. London, Sect. B **66**, 695–709 (1953).
[8] Y. A. Ilinskii, B. Lipkens, T. S. Lucas, T. W. Van Doren, and E. A. Zabolotskaya, "Nonlinear standing waves in an acoustical resonator," J. Acoust. Soc. Am. **104**, 2664–2674 (1998).
[9] C. Vanhille and C. Campos-Pozuelo, "Numerical model for nonlinear standing waves and weak shocks in thermoviscous fluids," J. Acoust. Soc. Am. **109**, 2660–2667 (2001).
[10] Y.-D. Chun and Y.-H. Kim, "Numerical analysis for nonlinear resonant oscillations of gas in axisymmetric closed tubes," J. Acoust. Soc. Am. **108**, 2765–2774 (2000).
[11] A. B. Coppens and J. V. Sanders, "Finite-amplitude standing waves within real cavities," J. Acoust. Soc. Am. **58**, 1133–1140 (1975).
[12] C. Campos-Pozuelo, B. Dubus, and J. A. Gallego-Juarez, "Finite-element analysis of the nonlinear propagation of high-intensity acoustic waves," J. Acoust. Soc. Am. **106**, 91–101 (1999).
[13] *Nonlinear Acoustics*, edited by M. F. Hamilton and D. T. Blackstock (Academic, New York, 1998).
[14] C. Vanhille and C. Campos-Pozuelo, "A high-order finite-difference algorithm for the analysis of standing acoustic waves of finite but moderate amplitude," J. Comput. Phys. **165**, 334–353 (2000).
[15] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (McGraw–Hill, New York, 1968).
[16] A. D. Pierce, *Acoustics* (Acoustical Society of America, New York, 1989).
[17] R. Bossart, N. Joly, and M. Bruneau, "Hybrid analytical and numerical solutions for acoustic boundary problems in thermo-viscous fluids," J. Sound Vib. **263**, 69–84 (2003).
[18] R. Bossut and J.-N. Decarpigny, "Finite element modeling of radiating

structures using dipolar damping elements,'' J. Acoust. Soc. Am. **86**, 1234–1244 (1989).
[19] B. Dubus, ''Coupling finite element and boundary element methods on a mixed solid-fluid/fluid-fluid boundary for radiation or scattering problems,'' J. Acoust. Soc. Am. **96**, 3792–3799 (1994).
[20] L. Elvira-Segura and C. Campos-Pozuelo, ''Experimental study of 3-D finite amplitude ultrasonic standing waves,'' *Nonlinear Acoustics at the Turn of the Millenium*, Proceedings of the 15th International Symposium on Nonlinear Acoustics, edited by W. Lauterborn and T. Kurz (American Institute of Physics, Melville, NY, 2000), pp. 177–180.
[21] J. Gallego-Juarez, G. Rodriguez-Corral, J. L. San Emeterio, and F. Montoya-Vitini, European Patent EP 450,030 (1991), US Patent 5,299,175 (1994).
[22] C. J. Moore, ''Measurement of radial and circumferencial modes in annular and circular ducts,'' J. Sound Vib. **62**, 235–256 (1979).
[23] D. B. Cruikshank, ''Experimental investigation of finite-amplitude acoustic oscillations in a closed tube,'' J. Acoust. Soc. Am. **52**, 1024–1036 (1972).

Nonlinear waveform distortion and shock formation in the near field of a continuous wave piston source

Oleg A. Sapozhnikov[a] and Vera A. Khokhlova
Department of Acoustics, Faculty of Physics, Moscow State University, Moscow 119992, Russia

Dominique Cathignol
INSERM, Unité 556, 151 Cours Albert Thomas, 69424 Lyon Cedex 03, France

(Received 27 April 2001; accepted for publication 15 February 2004)

A classical effect of nonlinear acoustics is that a plane sinusoidal acoustic wave propagating in a nonlinear medium transforms to a sawtooth wave with one shock per cycle. However, the waveform evolution can be quite different in the near field of a plane source due to diffraction. Previous numerical simulations of nonlinear acoustic waves in the near field of a circular piston source predict the development of two shocks per wave cycle [Khokhlova *et al.*, J. Acoust. Soc. Am. **110**, 95–108 (2001)]. Moreover, at some locations the peak pressure may be up to 4 times the source amplitude. The motivation of this work was to experimentally verify and further explain the phenomena of the nonlinear waveform distortion. Measurements were conducted in water with a 47-mm-diameter unfocused transducer, working at 1-MHz frequency. For pressure amplitudes higher than 0.5 MPa, two shocks per cycle were observed in the waveform beyond the last minimum of the fundamental harmonic amplitude. With the increase of the observation distance, these two shocks collided and formed one shock (per cycle), i.e., the waveform developed into the classical sawtooth wave. The experimental results were in a very good agreement with the modeling based on the Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1695433]

PACS numbers: 43.25.Jh, 43.28.Lv, 43.25.Vt [MFH]

Pages: 1982–1987

I. INTRODUCTION

A plane piezoelectric transducer can usually be considered as a piston source. In the limit of infinite aperture, it radiates a plane acoustic wave. If the aperture is finite, but large compared with the wavelength, the radiated wave may be considered as a quasiplane wave. Nonlinear evolution of plane acoustic waves is a well-studied phenomenon. It is interesting, therefore, to compare classical theoretical results for the plane waves with those obtained in the experiments with intense waves generated by piezoelectric sources of finite size. In the first experiments of this kind the transformation of an initially sinusoidal waveform into a sawtooth waveform was observed and it was concluded that the results were in a good accordance with the classical theory.[1–3] Later, the effect of diffraction on the nonlinear waveform was studied.[4–7] It was found that diffraction causes an asymmetry in the nonlinear waveform. However, the basic features of waveform distortion for a diffracted wave were found to be similar to those of a plane wave. In particular, the peak pressure during the shock formation is of about the same value as the initial wave amplitude, and only one shock is formed within each wave cycle. Most of the theoretical work was performed for Gaussian sources.

In an earlier paper,[8] the authors conducted a numerical study of the near field of an intense continuous wave (cw) piston source. Two theoretical models and corresponding numerical codes for the description of nonlinear acoustic beams were reported. In the first model, simulations were performed in the time domain. Diffraction, nonlinearity, and arbitrary frequency-dependent absorption were treated independently using the method of fractional steps with a second-order operator-splitting algorithm.[9] The second model was a newly developed frequency-domain algorithm based on the KZK equation.[8] The main feature of the second code developed was that it was optimized to reduce calculation time in the presence of shocks. The results of the two models were found to be in a very good agreement. The frequency-domain code was used to study the strongly nonlinear regime of beam propagation, that is, where shocks are developed close to the source. Somewhat unexpected features of the waveform distortion in the near field were predicted. Figure 1 shows typical on-axis waveforms at three increasing distances for the linear regime (dotted line) and nonlinear regime (solid line). It is seen that formation of two shocks per cycle is predicted, in contrast to the theory for plane waves. Another interesting feature is the amplification of the peak pressure: In Fig. 1(a) the peak pressure is 3 times larger than the source amplitude. Linear theory predicts a maximum is amplification of 2. The aforementioned effects are visible only within a limited range in the near field of the source. For larger distances the two shocks collide [Fig. 1(b)] and form one shock per cycle [Fig. 1(c)], i.e., the waveform develops into the usual sawtooth wave.

The motivation of this work was to experimentally verify and further explain the peak pressure amplification and existence of the ''double-shock'' regime.

[a]Electronic mail: oleg@acs366.phys.msu.ru

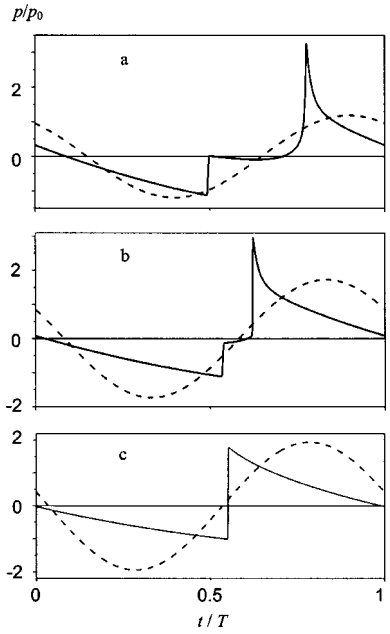


FIG. 1. Numerical predictions illustrating the formation of two shocks per cycle for waveforms on the axis of a circular piston with very high surface pressure. Dotted lines correspond to the waveform in the linear regime. The curves a, b, and c correspond to increasing distances from the source. The acoustic pressure, p, is normalized by initial wave amplitude, p_0; the time, t, is normalized by the wave period, T.

II. MATERIALS AND METHODS

A. Experimental setup

The experimental setup was similar to that used by Nachef *et al.*[10] and is shown in Fig. 2. The circular piston source had an aperture diameter of approximately 47 mm and an operating frequency of 1.0 MHz. The transducer was made from a PZT piezoceramics with an acoustic impedance of 31 MRayl (P762, Quartz et Silice, France). It had a matching quarter-wave plate of 3.4-MRayl impedance, at the front side, and a loaded epoxy wax backing of 4.5-MRayl impedance, at the rear side. Further details on the transducer assembly can be found in Ref. 10. The presence of the backing reduced parasitic plate waves, which appear in the piezoceramics in addition to the principal thickness vibration. If the necessary steps are not taken, these plate waves can significantly disturb the uniformity of the transducer surface vibration and, as a result, the source may not operate as an ideal piston.[11,12]

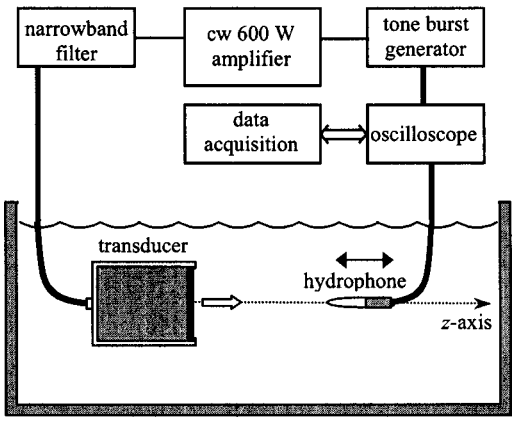


FIG. 2. Experimental setup.

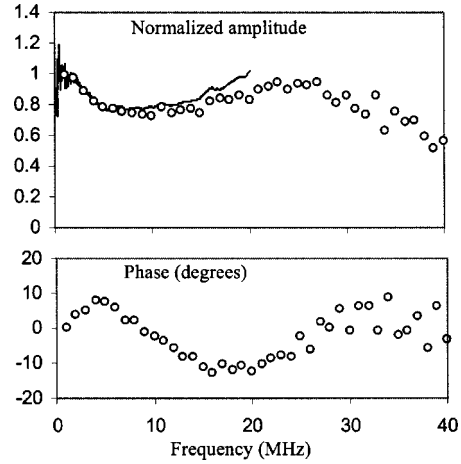


FIG. 3. Frequency response of the hydrophone. Top: amplitude sensitivity normalized by its value at 1 MHz. The solid curve represents the sensitivity (without preamplifier) provided by the hydrophone manufacturer (SEA); the circles show the measurements by the authors (with preamplifier). Bottom: phase shift as a function of frequency.

The source was placed in a $0.75\times0.6\times0.5$-m tank filled with degassed and demineralized water. To simulate a continuous wave (cw), the transducer was driven by a tone burst (30 cycles) from a waveform generator (8116A, Hewlett-Packard). The voltage from the generator was amplified by a 600-W power amplifier (Ampar 1004, Prana, France). To reduce nonlinear distortion of the electrical signal a narrowband filter, with central frequency of 1 MHz, was used after the power amplifier. The pressure field was measured with a broadband PVDF calibrated hydrophone with 0.3-mm active element diameter (GL049, SEA). The pressure waveform was recorded within a 10-μs window starting 20 μs after the first arrival of the signal. The signal from the hydrophone was amplified with a wide-band preamplifier (A17DB, SEA), digitized by an oscilloscope (Tektronix 2430A), and then transferred to a computer. An x–y–z computer-controlled positioning system was used to move the hydrophone along the axis of the transducer symmetry with an accuracy of 0.01 mm (Micro-Contrôle, France). The procedure of finding the axis of symmetry (acoustical axis) has been described previously.[10,12]

For correct representation of the nonlinearly distorted waveform, the hydrophone should have uniform amplitude and phase responses within the frequency range of the recorded signal. Although the hydrophone used was fairly broadband, its frequency response was not flat (Fig. 3). To avoid the corresponding waveform distortion, the recorded waveform was corrected based on the hydrophone frequency response. The latter was measured by dividing spectra of theoretically predicted and measured waveforms at a sufficiently large distance from the source, where the signal had sawtooth waveform with one shock per cycle. The hydrophone frequency response (amplitude and phase) is shown in Fig. 3. Circles represent the response to harmonics of 1 MHz; the corresponding values were used to correct the measured waveform. Solid line shows the manufacturer's

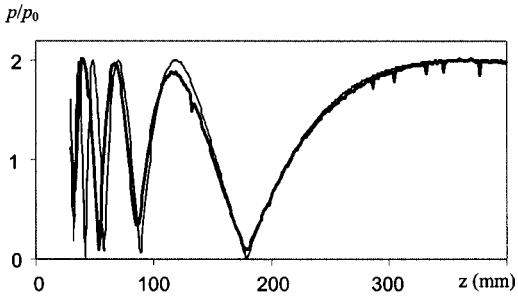


FIG. 4. Amplitude along the axis in the linear regime. The amplitude, p, is normalized by the source amplitude, p_0. Thick line represents measured amplitude; thin line is the theoretical curve for a piston of $a=23.2$-mm radius.

(SEA) data on the hydrophone calibration in the frequency band from 0.2 to 20 MHz. It is seen that our measurements (circles) coincide very well with this curve. Some decrease in the hydrophone sensitivity for high frequencies can be explained by the attenuation introduced to these frequencies by the preamplifier. The amplitude response in Fig. 3 is normalized by the hydrophone sensitivity at 1 MHz, which was given as 61.7 mV/MPa by a manufacturer. Note that the measured hydrophone response is somewhat irregular for the frequencies higher than 25 MHz; this is an indication of a possible poor accuracy. However, the corresponding frequency components in the signal spectrum were very small even for waves with shocks, i.e., their inaccurate detection could not significantly change the waveform.

B. Determination of the piston source parameters

A circular piston source can be characterized by three parameters: frequency, f, radius, a, and the normal velocity amplitude on the piston surface, u_n. Instead of u_n it is more convenient to use effective acoustic pressure amplitude of the source, $p_0=\rho c u_n$, where ρ and c are the ambient density and the sound speed, correspondingly. In the linear regime, an exact solution for the acoustic pressure distribution along the axis is $p=2p_0|\sin[(\pi f/c)(\sqrt{a^2+z^2}-z)]|$, where p is the amplitude and z is the axial distance from the source.[13]

The experimental values for the frequency and the sound speed were $f=1$ MHz and $c=1490$ m/s. The value of the effective source radius, a, was determined by comparing the measurements of the axial distribution of the acoustic pressure in the linear regime with the theoretical solution. Figure 4 compares the axial distribution of the peak-to-peak value of the hydrophone signal normalized by the amplitude of the last (outermost) lobe, and the theoretical distribution for a piston of radius $a=23.2$ mm. The choice of a gives the same coordinate of the last minimum as that obtained experimentally. It is seen that the two curves coincide very well for distances larger than 60 mm from the source. At smaller distances discrepancies arise due to the finite size of the hydrophone and because the transducer velocity distribution is not perfectly uniform due to existence of plate waves in the piezoceramics.[11,12] A radius of $a=23.2$ mm was used in the numerical modeling.

Another parameter necessary for the modeling of a nonlinear field is the source pressure amplitude p_0. In the linear regime, the value of p_0 can be found as half of the amplitude of the last lobe (see Fig. 4). In the nonlinear regime, the amplitude of the last lobe is decreased due to nonlinear attenuation, so p_0 cannot be measured in the same way. To avoid this difficulty, the wave amplitude was measured close to the source before nonlinear effects become pronounced. The waveforms in the linear and nonlinear regimes are shown in Fig. 5. The two signals on the top are for a low transducer voltage ($U_{\mathrm{pp}}=13$ V peak-to-peak); the two signals at the bottom are for the maximum voltage used ($U_{\mathrm{pp}}=516$ V peak-to-peak.) The right-hand side plots correspond to the position of the last lobe maximum, $z=370$ mm. In the low-voltage case the waveform at this distance is sinusoidal, as would be expected in the linear one. Therefore, given the hydrophone sensitivity and the linear solution for the on-axis pressure, one can determine the transducer pressure amplitude p_0. In contrast, at the high transducer voltage, the waveform is strongly distorted due to acoustic nonlinearity, and linear theory cannot be used to determine p_0. The signals on the left-hand side correspond to measurements at $z=40$ mm. It is seen that, even at the maximum level of the transducer voltage, the waveform remains sinusoidal, i.e., nonlinear acoustic effects can be neglected here. Therefore, the ratio between the signal amplitudes at the two levels can be considered equal to the corresponding ratio for p_0 and the source pressure in the nonlinear regime can be determined. These measurements were repeated at different levels of the transducer voltage, and it was found that the value of p_0 varied linearly with the transducer voltage. The maximum voltage $U_{\mathrm{pp}}=516$ V corresponded to a pressure amplitude $p_0=1.01$ MPa.

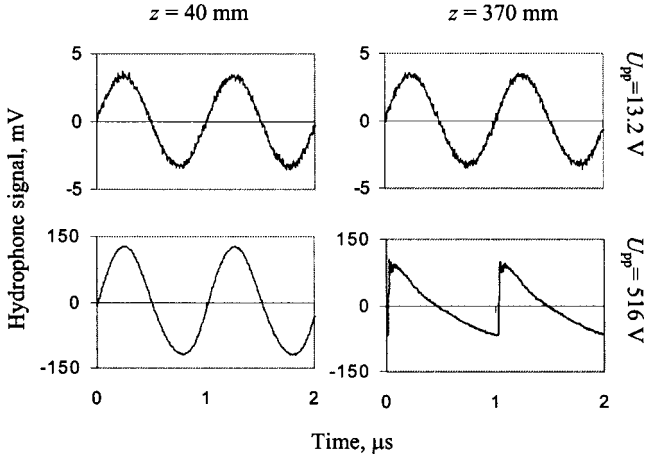


FIG. 5. Measured waveforms at $z=40$ and 370 mm, for two levels of source peak-to-peak voltage $U_{\mathrm{pp}}=13.2$ and 516 V.

C. Theory

The nonlinear acoustic field radiated by a circular piston source can be modeled by the Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation.[4] The KZK equation can be written in terms of the axial component of the particle velocity u as

$$\frac{\partial}{\partial\tau}\left(\frac{\partial u}{\partial z}-\frac{\beta}{c^2}u\frac{\partial u}{\partial\tau}-\frac{b}{2c^3\rho}\frac{\partial^2 u}{\partial\tau^2}\right)=\frac{c}{2}\Delta_{\perp}u. \tag{1}$$

Here, $\tau = t - z/c$ is the retarded time, ρ is the ambient density, β is the nonlinear parameter, and b is the dissipative parameter of the medium, $\Delta_\perp = \partial^2/\partial r^2 + r^{-1}\partial/\partial r$ is the transverse Laplacian, and r is the lateral coordinate (distance from the axis of symmetry). Within the framework of the KZK equation it is valid to use the plane-wave impedance relationship $p = \rho c u$ to transform between acoustic pressure p and particle velocity u. The boundary condition at $z = 0$ appropriate for a uniform baffled piston is: $u = (p_0/\rho c) \cdot \sin(2\pi f t)$ if $r \leqslant a$ and $u = 0$ if $r > a$. The water parameters are $\rho = 10^3$ kg/m^3, $\beta = 3.5$, b $= 4.33 \times 10^{-3}$ kg/(s·m), $c = 1.49 \times 10^3$ m/s, the source parameters are $f = 1$ MHz, $a = 2.32 \times 10^{-2}$ m, p_0 varies from 0.0258 to 1.01 MPa. Equation (1) is solved numerically using a frequency-domain code described previously.[8]

III. RESULTS AND DISCUSSION

A. Comparison of experimental data and numerical results

The measurements and corresponding modeling were performed for various levels of the source pressure amplitude p_0. For pressure amplitudes higher than 0.5 MPa, two shocks per cycle were observed in the waveform beyond the last minimum of the fundamental harmonic amplitude. However, results are presented only for the maximum source amplitude $p_0 = 1.01$ MPa. Nonlinear waveform distortion and shock formation were most pronounced in this case. Figure 6 shows theoretical and experimental waveforms at various distances from the source. The waveforms are shown within a 2-μs temporal window, i.e., for two cycles of the radiating frequency 1 MHz. To illustrate the complexity of the near-field structure, the simulated 2D distribution of the wave amplitude in linear regime is shown in the middle of the picture. Circles mark the locations where the waveforms were measured. It is seen from Fig. 6 that all the experimental curves are nearly identical to the corresponding theoretical waveforms. At relatively small distance $z = 70$ mm, the waveform is distorted similar to the case of a plane wave. At larger distances the distortion becomes unusual. For instance, at $z = 170$ mm the waveform is shocked, but, contrary to the case of a plane wave, the waveform includes a local maximum, i.e., it is not monotonic between the shocks. This can be explained by the fact that at $z = 170$ mm in the linear regime the wave amplitude is close to zero (see the 2D beam structure in Fig. 6.) As a result, the second harmonic predominates over the fundamental and the waveform appears to double its frequency. For longer distances, the local maximum between the shocks increases (see, e.g., $z = 210$ mm), giving rise to a shock formation at $z = 250$ mm. At this point there are two shocks per wave cycle. Further propagation of the wave leads to collision of these two shocks and only one shock remains ($z = 280$ and 320 mm).

Shown in Fig. 7 are experimental waveforms that illustrate in more detail the formation of the second shock and interaction of the two shocks. Note that only one cycle is presented. The first shock formed at approximately $z = 130$ mm (data not shown). At the first shown waveform ($z = 220$ mm) this shock sits in the rarefaction phase of the

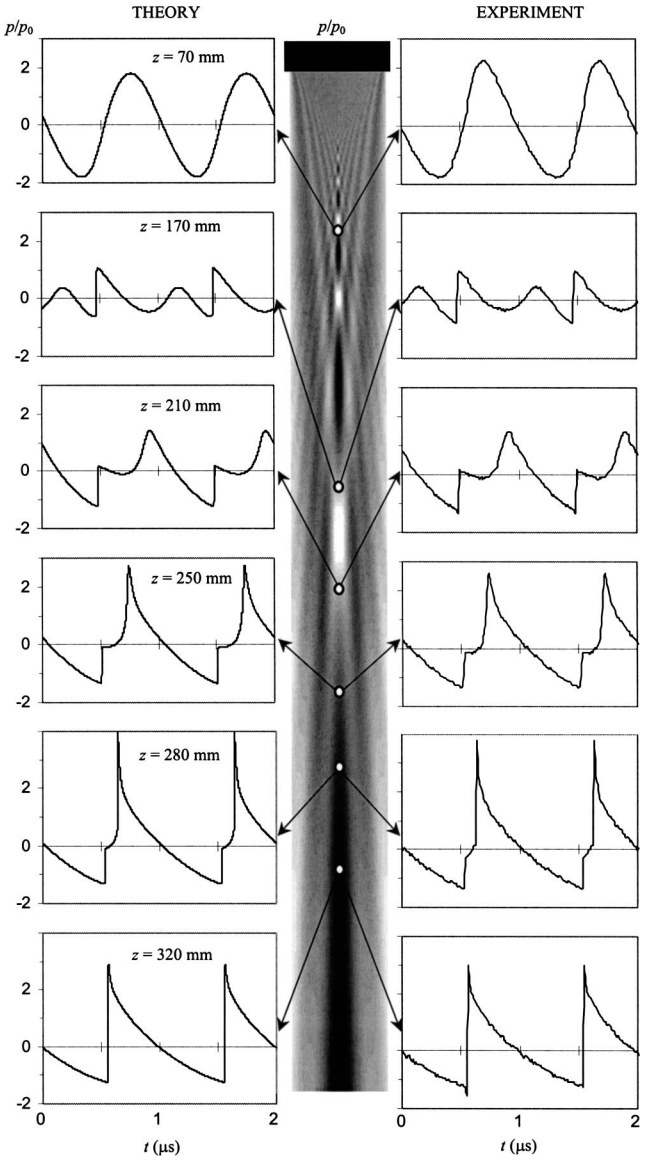


FIG. 6. Comparison of the theoretical (left-hand side) and the experimental (right-hand side) waveforms on the acoustic axis at various distances z from the source. Drawn in the middle is the theoretical 2D distribution of the wave amplitude in linear regime (linear gray scale). Darker regions correspond to higher amplitudes.

waveform. Because of that, it propagates slower than the low-amplitude signal, i.e., moves from left to right in the waveform curves of Fig. 7. A peak in the following compression phase increases in amplitude with distance and finally also becomes shocked ($z = 260$ mm). It propagates faster than the low-amplitude signal, i.e., moves from right to left in Fig. 7 ($z = 260$–300 mm). This behavior is in a qualitative agreement with the properties of plane nonlinear acoustic waves.[5] Since the second shock moves faster than the first one ($z = 260$–300 mm), finally they collide ($z = 320$ mm). After the collision the waveform contains only one shock per cycle. This sawtooth wave has an asymmetric shape: the compression phase is shorter than the rarefaction one and the positive peak pressure exceeds the negative peak pressure. As a result, the shock front propagates faster than the low-amplitude signal ($z = 320$ and 340 mm). This kind of waveform is typical for the diffracted sound beams.[4]

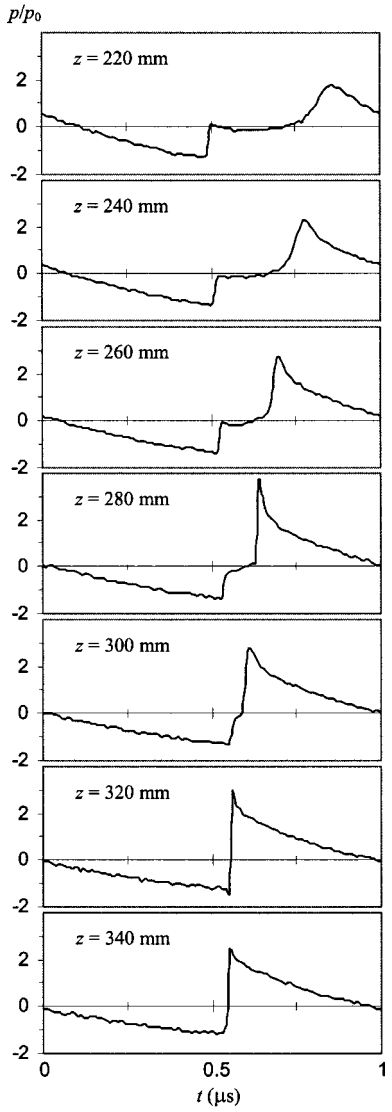


FIG. 7. Experimental waveforms showing the interaction of the two shocks.

B. Discussion

The observed nonlinear evolution of the acoustic wave radiated by a piston source is quite unusual compared to the accepted nonlinear behavior of a plane wave or a wave from a Gaussian source. The phenomenon observed here can be explained in terms of direct and edge waves. The decomposition of the wave into these two waves is frequently used in the analysis of the transient radiation from a piston.[14] The direct wave is a plane wave propagating along the transducer axis. The edge wave comes from the transducer rim and has a toroidal wavefront. In the nonlinear case the principle of superposition is no longer valid, so the direct and the edge wave cannot be considered as independent. On the other hand, in the near field they are noncollinear and, as a result, interact weakly. Furthermore, the edge wave has much smaller amplitude on the path from the edge to the axis. It is only near the axis that the amplitude of the edge wave is close to that of direct wave, due to simultaneously arriving signals from the entire transducer rim. The wave radiated by the piston source, therefore, can be approximated as a superposition of nonlinear plane wave and linear edge wave

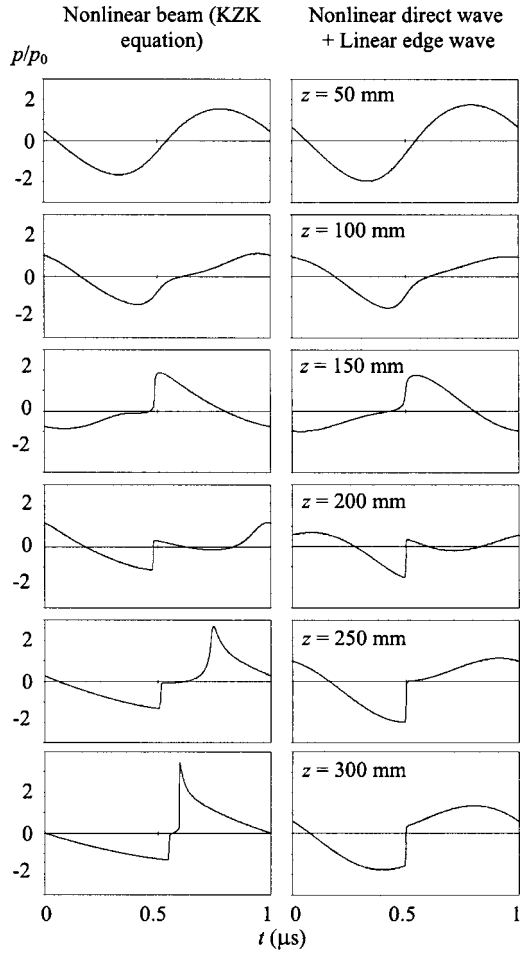


FIG. 8. Waveform evolution calculated from different models: The left side corresponds to the KZK equation and the right side corresponds to Eqs. (2)–(4). There is excellent agreement for $z<250$ mm.

$$p = p_D + p_E . \tag{2}$$

Similar interpretation of experimentally observed nonlinear wave distortion was previously used by Hobaek and Ystad in the study of a focused beam.[15] They considered the edge wave as a linear ''background signal'' for the shocked direct wave. However, they did not observe nor predict a second shock associated with the edge wave.

The nonlinear evolution of a plane wave (direct wave) is governed by the Burgers equation

$$\frac{\partial p_D}{\partial z} - \frac{\beta}{c^3 \rho} p_D \frac{\partial p_D}{\partial \tau} = \frac{b}{2c^3 \rho} \frac{\partial^2 p_D}{\partial \tau^2}, \tag{3}$$

with the boundary condition $p_D = p_0 \cdot \sin(2\pi f \tau)$ at $z=0$. Equation (3) was solved numerically using a frequency-domain finite difference algorithm.[8] The edge wave on the axis is an inverted replica of the source waveform, retarded in accordance with the distance to the edge

$$p_E = -p_0 \cdot \sin\left[2\pi f \left(\tau - \frac{\sqrt{z^2+a^2}-z}{c} \right) \right]. \tag{4}$$

Figure 8 compares the waveforms calculated using the KZK equation (1) and those from Eqs. (2)–(4). It is seen that the waveforms of the two models are almost indistinguishable at small distances ($z \leqslant 150$ mm), even when the wave-

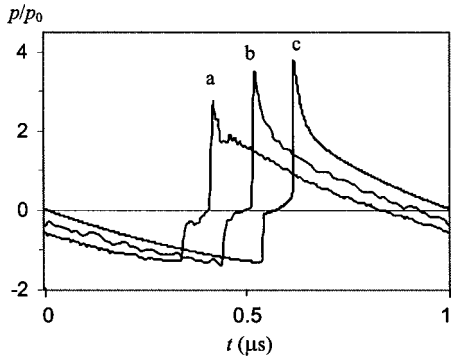


FIG. 9. Observation of nonlinear increase of the peak pressure. (a) Signal at the output of the hydrophone; (b) the hydrophone signal corrected using the frequency response; (c) theoretical waveform. The waveforms *a* and *b* are shifted for easier comparison.

form is shocked. This confirms the behavioral assumptions used for the direct and edge waves. For larger distances, the nonlinearity influences the edge wave as well and discrepancies between results of the two models appear, especially in the structure of the compression part of the waveform (Fig. 8). Growth of the positive peak pressure can be interpreted as a result of nonlinear focusing of the edge wave near the axis. Further propagation of the wave can no longer be considered as a superposition of two independent waves. The wave behaves more like a nonlinear plane wave. The negative pressure shock propagates slower than the linear wave; in contrast, the positive pressure shock propagates faster. This results in collision of the two shocks and formation of one shock.

In addition to the double-shock structure of the waveform, our earlier modeling[8] predicted an increase of the positive peak pressure (3–4 times) compared to the source pressure p_0. This increase can be interpreted as a consequence of the edge wave nonlinear focusing. Indeed, it is known that nonlinearity may result in better focusing, because the higher harmonics are less diffractive.[4] As it was shown previously,[8] the effect of the peak pressure amplification is highly localized both in time (0.05–0.1 μs) (see the waveform at $z = 280$ mm in Fig. 6) and in space in the lateral direction (0.5 mm). Therefore, experimental observation of the peak pressure increase is possible only by a hydrophone with sufficiently good response and with sufficiently small active element. Also, a very accurate positioning of the hydrophone on the axis of symmetry is needed. The importance of the hydrophone response is illustrated in Fig. 9. The hydrophone signal before calibration (a) shows a lower peak than the waveform with the frequency response correction (b). The theoretical waveform is shown by curve (c) and is in good agreement with the corrected waveform. The experimental curves are shifted in time by 0.4 μs (a) and 0.2 μs (b), correspondingly, for easier comparison. It is seen that more than 3 times increase in the peak is not observed without the frequency response correction. The fact that the peak pressure can be up to 100% larger than the peak predicted by linear theory is an important observation that is especially relevant at ultrasound intensity levels that could cause damage to the propagation medium.

IV. CONCLUSIONS

Nonlinear distortion of the pressure waveform in water in the near field of an intense circular piston source was investigated. The experimental measurements confirmed the theoretical prediction of amplification of peak pressure due to nonlinear diffraction, the formation of two shocks per wave cycle, and their subsequent collision. This nonlinear evolution of a wave in the near field of a piston is very different from that of a plane wave and can be interpreted in terms of direct and edge waves. The results of the measurements are in a good agreement with theory based on the KZK equation. The effects studied in the present paper have not been observed before; however, they may be of importance for various applications of piston ultrasound sources operating at high levels, e.g., therapeutic application of intense ultrasound.

ACKNOWLEDGMENTS

This work was partly supported by "L'action Spécifique pour les Pays d'Europe Centrale et Orientale (PECO) de l'Institut National de la Santé Et de la Recherche Médicale (INSERM)," the Russian Foundation for the Basic Research (RFBR), and CRDF. The authors are grateful to Dr. Robin Cleveland for discussions and proofreading the text.

[1] V. A. Burov and V. A. Krasil'nikov, "Direct observation of the intensive ultrasonic wave sharp distortion in a liquid," Dokl. Akad. Nauk SSSR **118**, 920–923 (1958).
[2] K. A. Naugol'nykh and E. V. Romanenko, "On the problem of finite-amplitude wave propagation in liquids" (in Russian), Akust. Zh. **4**, 200–202 (1958).
[3] I. G. Mikhailov and V. A. Shutilov, "Finite amplitude ultrasonic wave shape distortion in various liquids" (in Russian), Akust. Zh. **6**, 340–346 (1960).
[4] N. S. Bakhvalov, Ya. M. Zhileikin, and E. A. Zabolotskaya, *Nonlinear Theory of Sound Beams* (American Institute of Physics, New York, 1987).
[5] O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics* (Plenum, New York, 1977).
[6] V. G. Andreev, A. A. Karabutov, and O. V. Rudenko, "Experimental investigation of finite amplitude sound beams," Moscow Univ. Phys. Bull. **39**, 35–38 (1984).
[7] D. R. Bacon, "Finite amplitude distortion of the pulsed fields used in diagnostic ultrasound," Ultrasound Med. Biol. **10**, 189–195 (1984).
[8] V. A. Khokhlova, R. Souchon, J. Tavakkoli, O. A. Sapoznikhov, and D. Cathignol, "Numerical modeling of finite amplitude sound beams: Shock formation in the near field of a cw plane piston source," J. Acoust. Soc. Am. **110**, 95–108 (2001).
[9] J. Tavakkoli, D. Cathignol, R. Souchon, and O. A. Sapozhnikov, "Modeling of pulsed finite-amplitude focused sound beams in time domain," J. Acoust. Soc. Am. **104**, 2061–2072 (1998).
[10] S. Nachef, D. Cathignol, J. N. Tjotta, A. M. Berg, and S. Tjotta, "Investigation of a high intensity sound beam from a plane transducer. Experimental and theoretical results," J. Acoust. Soc. Am. **98**, 2303–2323 (1995).
[11] B. Delannoy, C. Bruneel, F. Haine, and R. Torguet, "Anomalous behavior in the radiation pattern of piezoelectric transducers induced by parasitic Lamb wave generation," J. Appl. Phys. **51**, 3942–3948 (1980).
[12] D. Cathignol, O. A. Sapozhnikov, and J. Zhang, "Lamb waves in piezoelectric focused radiator as a reason for discrepancy between O'Neil's formula and experiment," J. Acoust. Soc. Am. **101**, 1286–1297 (1997).
[13] A. D. Pierce, *Acoustics* (Acoust. Soc. Am., New York, 1994), pp. 231–233.
[14] D. T. Blackstock, *Fundamentals of Physical Acoustics* (Wiley, New York, 2000), pp. 460–461.
[15] H. Hobaek and B. Ystad, "Experimental and numerical investigation of shock wave propagation in the post focal region of a focused sound field," Acust. Acta Acust. **83**, 978–986 (1997).

Passive amplification and directivity from air-coupled surface waves generated above a structured ground

Gilles A. Daigle[a] and Michael R. Stinson
Institute for Microstructural Sciences, National Research Council, Ottawa, Ontario K1A 0R6, Canada

(Received 1 July 2003; revised 7 February 2004; accepted 17 February 2004)

An outdoor experiment is reported in which air-coupled surface waves are generated in the frequency range between 63 and 130 Hz above a strip of structured ground of finite width. Since the sound pressure above the structured ground can be greater than if the ground was rigid, passive amplification can be obtained. The experimental results are compared to theoretical calculations that use a boundary element method. Both predict gains of up to 6 dB, relative to the pressure on a rigid plane when the strip width is of the order of one wavelength. Due to their finite width, the surface wave strips have a directional response. For a source at a large distance from the strip, the sound pressure level is up to 6 dB greater for sound arriving along the axis of the strip than transverse. For sources near the strip, as might arise in a near-field beamforming application, this difference can be as much as 25 dB. [DOI: 10.1121/1.1703535]

PACS numbers: 43.28.En, 43.20.Mv, 43.20.El, 43.50.Vt [LLT] Pages: 1988–1992

I. INTRODUCTION

Directional microphones coupled to parabolic reflectors are commonly used for the detection and monitoring of higher-frequency sounds. However, in the case of low-frequency sounds, a parabolic reflector would become expensive and unwieldy due to the longer wavelengths. Thus, for example, Benson *et al.*[1] have described a simple baffle-type directional microphone that can be used as a fixed installation for low frequency sound monitoring. On the other hand, Brekhovskikh[2] has discussed how air-coupled surface waves could be used for sound wave pickup with high directivities associated with small transverse dimensions.

A series of controlled laboratory experiments have been recently described[3] in which air-coupled surface waves are generated in the frequency range between 800 and 1700 Hz above a surface composed of a lattice of small cavities. In this paper we show that the sound pressure above the lattice can be greater than if the surface was rigid and passive amplification is obtained. Moreover, if the lattice of cavities is restricted to a strip of finite width embedded in an otherwise rigid surface, amplification will depend on the direction of the incident sound and directional receivers can be designed.

In this paper, we discuss the construction of a structured ground of finite width to obtain passive amplification and directivity outdoors at frequencies around 100 Hz. Experiments were carried out above the structured ground and the measurements are compared with numerical results obtained using a boundary element method. It will be shown that a directional receiver can be designed that provides an alternative to beamforming or other techniques for the long-range detection of acoustic signatures at low frequencies. Applications in the area of noise control will also be discussed.

II. THEORY AND NUMERICAL CALCULATIONS

Numerical calculations were performed using a boundary element method (BEM). A commercial software package SYSNOISE (version 5.5) was used for this purpose, running on a personal computer. Since the study of surface waves that propagate slower than the free-space speed of sound is not a typical application for this software, a series of test calculations was discussed in Ref. 3 to establish the applicability and accuracy of the computational software. It was found that SYSNOISE can be applied with confidence to the situation of interest, a strip of finite impedance, as shown in Fig. 1.

There are several modules in SYSNOISE that permit the kinds of calculations that will be described here. It was determined that the best compromise between speed of computation, flexibility, and accuracy was achieved using the ''direct (collocation)'' BEM option. To implement this module, a planar mesh of elements in the $z=0$ plane (ground surface) is used to represent the rectangular region of interest and a plane of symmetry invoked in the $z=0$ plane to effectively place the mesh in an infinite rigid baffle. To ensure accuracy in the calculations, mesh elements had linear dimensions less than 1/8 the wavelength of sound (free field) for the highest frequencies of interest. Where possible, use was made of symmetry to reduce the computational load.

III. CONSTRUCTION OF THE STRUCTURED GROUND

The general conditions for the formation of air-coupled surface waves are discussed in detail in Ref. 4. In particular, surface waves can be formed above cavities of finite depth (a lattice of cavities[3,4] or a comb-like surface[2]). The impedance of the cavities, and thus the formation of the surface waves, is controlled in large part by the depth of the cavities. In the case of an idealized surface with no resistance, the impedance is simply given by[2,4]

$$Z/\rho c = i\cot(kl), \tag{1}$$

for $\exp(-i\omega t)$ time convention, where l is the depth of the cavities and $k=2\pi/\lambda$, where λ is the wavelength. In the case of the previous laboratory experiments, the depth of the cavities was 2.56 cm and the frequencies of interest were

[a]Corresponding author. Electronic mail: gilles.daigle@nrc-cnrc.gc.ca

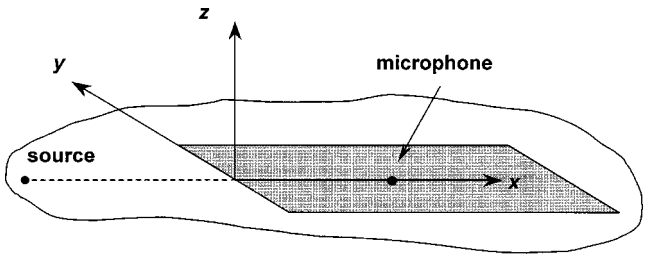


FIG. 1. Sketch showing a strip of finite impedance (the shaded area) in the x-y plane embedded in an otherwise hard ground.

around 1200 ± 400 Hz. In this paper, the frequencies of interest are around 100 ± 40 Hz and, thus, the depth of the cavities must be increased by a factor slightly greater than 10, or >25.6 cm.

The finished ground structure is shown in Fig. 2. It consists of a strip of comb-like surface constructed by first digging a trench-of the desired width—of the order of one wavelength (see Ref. 3). The length of a 4 ft$\times$8 ft sheet of commercial plywood is 244 cm. For the purpose here, this width was considered sufficiently close to the wavelengths of interest and plywood sheets were used to construct the comb-like surface. Sheets of 1/4 in.-plywood were cut lengthwise into strips of 29.2 cm by 244 cm (thus, four strips could be produced per sheet of plywood). The trench was filled with stone dust to the desired depth (29.2 cm). Sheets of 3/4 in.-plywood were also cut and used for the borders parallel to the strip. Grooves were cut into the borders every 10.2 cm and the strips were inserted into the grooves of two opposing borders and pressed into the stone dust to form the cavities. The relevant dimensional quantities are shown in Fig. 3 (not to scale) and the final dimensions of the comb-like structure were $l=27$ cm, $w=244$ cm, $a=9.5$ cm, and $b=10.8$ cm. The total length of the strip was 46.5 m.

In practice, Eq. (1) does not adequately describe the impedance of a realistic ground structure. The finite thickness of the strips introduces an effective porosity and leaks introduce a finite resistance.[3] The following equation was assumed to better describe the impedance of the ground structure:

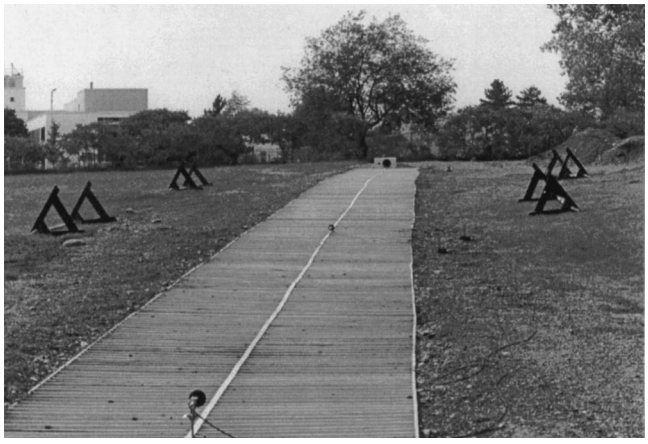


FIG. 2. Picture showing the strip of structured ground. The source can be seen in the distance as well as some of the measurement microphones positioned in the center of the strip. The width of the strip was 2.44 m and its total length was 46.5 m.

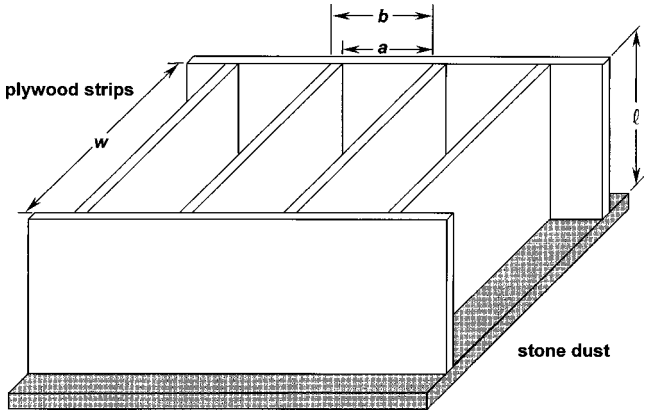


FIG. 3. The relevant dimensions of the comb-like structure (not to scale): $l=27$ cm, $w=244$ cm, $a=9.5$ cm, and $b=10.8$ cm. The top of the structure is flush with the ground surface.

$$Z=R(f)-i\left(\frac{b}{a}\right)\rho c\cot(kl), \tag{2}$$

where f is the frequency. The real part of Eq. (2) was determined by a series of measurements, as described in the next section. Note that Eq. (2) is strictly appropriate only for propagation in the x direction.

IV. DETERMINATION OF THE SURFACE RESISTANCE

A series of measurements was first made to fully characterize the impedance of the strip. A source was located on the ground in front of the strip at $x_s=-0.5$ (see Fig. 2). The source was the woofer unit of a commercial hi-fi loudspeaker. The center of the woofer was at a height $z_s=0.16$ m. A series of microphones were mounted along the centerline of the strip at a height of 0.1 m. The microphones were piezoelectric microphone-preamplifier units with 10 cm foam wind screens. The electrical signal from each microphone was fed to a sound level meter. Each microphone was calibrated using the manufacturer's calibrator before making the measurements.

Measurements were made on two different days in early fall. On the first day, the air temperature was 22 °C with light winds and the strip was cross-wind (cloud cover was not recorded). The second day was mainly sunny with an air temperature of 10 °C and light winds. The sound source was upwind. The two sets of points in Fig. 4 show the measured levels at 63, 100, and 130 Hz for two days. There is very good agreement between the two measured datasets and it was assumed that meteorological effects could be neglected.

Predictions corresponding to this geometry were then obtained using SYSNOISE. Calculations were repeated at each frequency, varying the value of R in Eq. (2), until the best fit to the measured data was obtained. The solid curves in Fig. 4 are the final results of this data matching, showing good agreement. The values of R thus obtained for all the frequencies used during the measurements are plotted in Fig. 5, and include the estimated uncertainty for best fit. The curve drawn through the points is obtained by an empirical

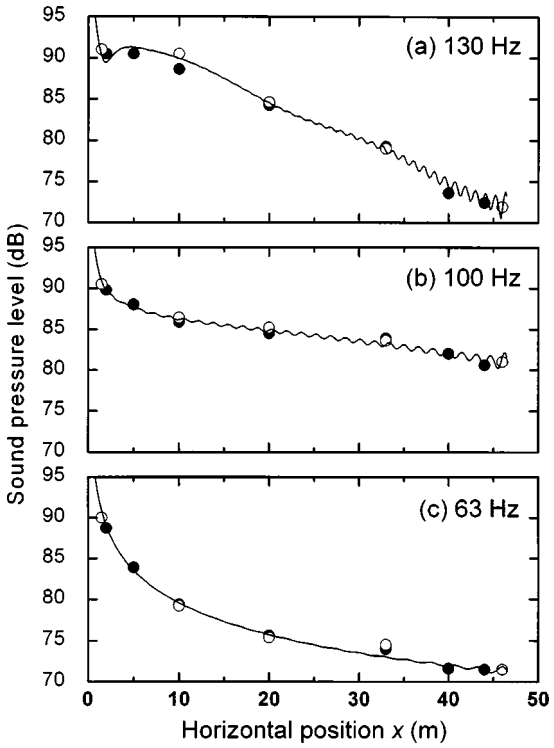


FIG. 4. The points are sound pressure levels measured as a function of horizontal position x along the center of the strip. The open and solid circles were obtained on two different days (22 °C with a light cross-wind and mainly sunny 10 °C downwind, respectively). The source is located at position $x_s = -0.5$ m. The curves are the corresponding BEM calculations obtained by adjusting R in Eq. (2) for best fit.

fitting procedure. The measurements show a rapid variation between 70 and 100 Hz. Thus, it was found necessary to use the following functional form:

$$R(f) = \frac{192 + 1220(f/100)^8}{1 + 16.65(f/100)^{10}}, \tag{3}$$

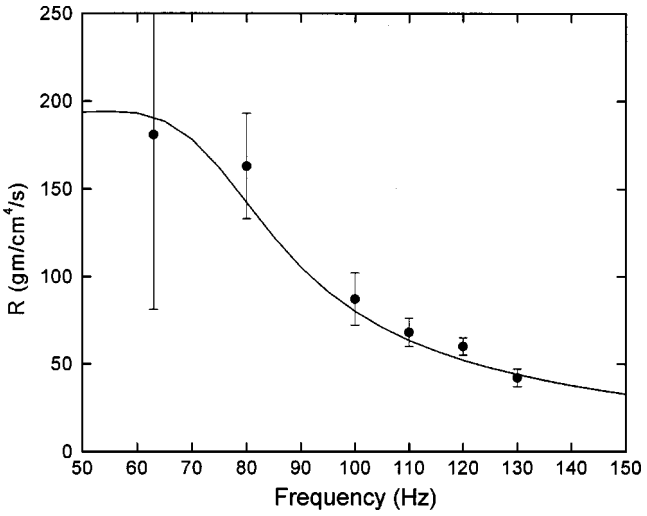


FIG. 5. The points are the real part of the impedance, R in Eq. (2), determined from measured levels above the ground structure (see Fig. 4). The curve is Eq. (3) and provides the real part of Eq. (2) when predicting amplification and directivity of the ground structure using BEM calculations.

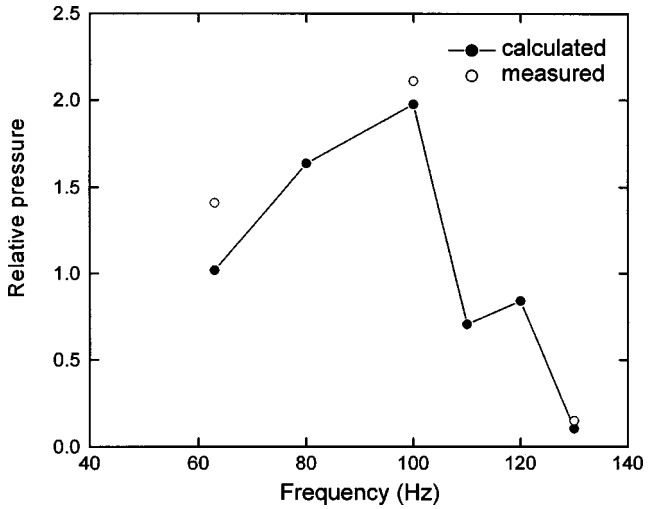


FIG. 6. A comparison between BEM calculations and measurements above the ground structure. The source is located at position $x_s = -31$ m and the microphone is located at position $x = 40$ m. The pressure is relative to the pressure at a distance of 31 m from the source above a hard surface (at $x = 0$, but in the absence of the ground structure).

with R having units of g/cm^4/s to describe the data. Equation (3) together with Eq. (2) was used for calculations in the next section.

V. AMPLIFICATION AND DIRECTIVITY

In order to measure the amplification due to the formation of surface waves above the structured ground, the source was moved back to position $x_s = -31$ m. A reference microphone was located at position $x = -3$ m. The ground between the source and the reference microphone was fine quarry dust, very hard packed by vehicles. Measurements confirmed that this ground had an effective flow resistivity[5] high enough to be considered rigid at the frequencies of interest here. Thus, the levels measured by the reference microphone were corrected using inverse square law to obtain the pressure at the beginning of the strip ($x = 0$), but in the absence of the strip. As shown in Ref. 3, this reference pressure is chosen to illustrate that the surface wave initially grows with increasing horizontal position ($x > 0$), reaching pressures that exceed the pressures at $x = 0$ above the hard surface.

The laboratory measurements reported in Ref. 3 showed that the maximum gains at 1200 Hz were obtained at horizontal positions between 1 and 3 m (or between 4 and 10 wavelengths) when the strip width was of the order of one wavelength. Thus, one can expect the maximum gains at 100 Hz above the outdoor ground structure at horizontal positions between about 15–35 m. A comparison of the data in Fig. 4(b) with the decrease in sound pressure level expected in the absence of the ground structured—i.e., compared to inverse square law—reveals that we might expect maximum gains between 25–45 m. Thus, for the purpose here, a microphone was located at horizontal position $x = 40$ m.

The open circles in Fig. 6 are the relative levels measured at $x = 40$ m above the strip. At 100 Hz, the measurements show a doubling of pressure providing a gain (ampli-

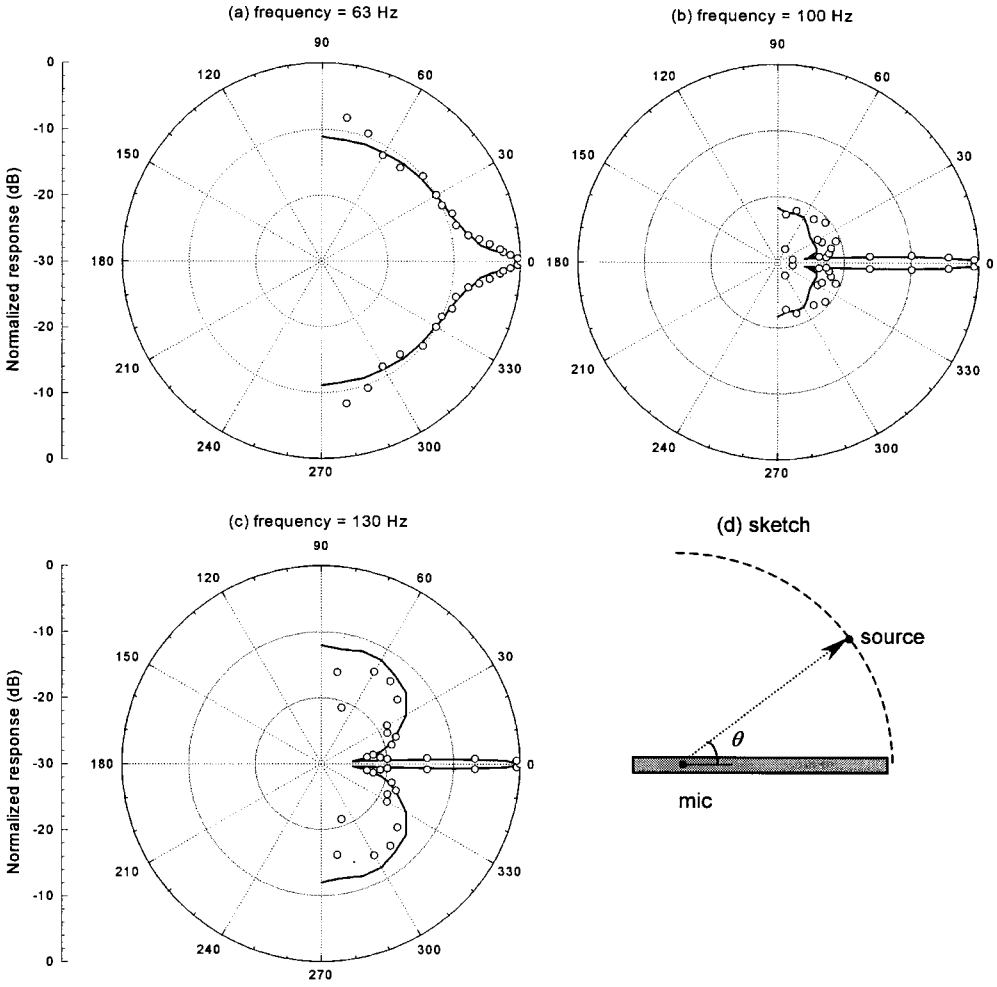


FIG. 7. A comparison between BEM calculations (curves) and measured (points) directional responses of the ground structure in the near field at frequencies of 63, 100, and 130 Hz. The receiver is located at position $x=40$ m. The angle of 0° corresponds to the source pointing at the strip ($y=0$) along the x axis and $x_s=-0.5$ m. The measurements were obtained by rotating the source about the microphone position, in the x-y plane, keeping the source–microphone distance constant, as shown in the sketch. The pressure is normalized to give 0 dB at the maximum and the circles are spaced every 10 dB.

fication) of 6 dB. The solid circles are the results predicted using SYSNOISE. A gain of 6 dB at 100 Hz above the outdoor ground structure is in agreement with the gains seen during the laboratory experiments at higher frequencies.[3]

In Ref. 3, other strip geometries and modifications that resulted in additional gains were discussed and could be applied to the outdoor ground structure. We also note that Albert[6] has recently reported evidence for the formation of the air-coupled surface wave above snow layers of finite depth over frozen ground. Thus, it would be conceivable to construct a temporary *ad-hoc* ground structure from a strip of snow layer of finite depth.

In addition to amplification, the ground structure has a directional response. Since the structured ground is restricted to a strip of finite width, the amplification will depend on the direction of the incident sound. Sound arriving transverse to the strip (i.e., in the y direction) does not develop surface waves since propagation is parallel to the comb-like structure. Thus there is a directional advantage of up to 6 dB at 100 Hz.

Significantly sharper directional responses are expected when the source is near the strip (i.e., in the near field). This is illustrated by the measurements shown in Fig. 7. These measurements were obtained by rotating the source about the microphone position, in the x-y plane, keeping the source–microphone distance (40.5 m) constant (as shown in the sketch). The angle of 0° corresponds to the source pointing at the strip ($y=0$) along the x axis and $x_s=-0.5$ m, just in front of the strip. The levels have been normalized to give 0 dB at the maximum of the main lobe (at 0°). The points in Fig. 7 are the measured directional responses of the ground structure at 63, 100, and 130 Hz (results were only measured for $0\leqslant\theta\leqslant 90°$ and they are mirrored for $270°\leqslant\theta\leqslant 360°$). The solid curves are the predicted directional responses using SYSNOISE [we are assuming that Eqs. (2) and (3) can be used for the source off the x-axis, even though, strictly, they are not applicable]. The results show a sharp narrow lobe at the two higher frequencies, with the center (0°) of the main lobe showing levels reaching an average of about 25 dB higher at 100 Hz relative to the signals off the main lobe (at angles $>5°$). The directional responses shown in Fig. 7 are very similar to the ones obtained during the laboratory experiments,[3] but here the outdoor measurements are confirmed by prediction (no prediction was available in the case of the laboratory experiments).

VI. CONCLUSION

A ground structure was constructed and was shown to behave as a directional receiver with passive amplification. The ground structure is an impedance strip of finite width and air-coupled surface waves are generated at frequencies where the surface impedance is a *spring-like* reactance and is greater than the resistive component. The width of the ground structure is tuned to obtain maximum amplification (gain) at frequencies around 100 Hz. Gains of about 6 dB have been measured and predicted near 100 Hz. This detector provides, for example, an alternative to beamforming or other techniques for the long-range detection of acoustic signatures at low frequencies.

In addition, since the amplification depends on the direction of the incident sound, the ground structure behaves as a directional receiver. When the source is close to the ground structure, the directional responses show sharp narrow lobes along the direction of the strip. At 100 Hz, the measured and predicted levels along the center of the main lobe are an average of about 25 dB higher than the levels off the main lobe. This directional characteristic could, for example, be exploited to distinguish the signal of one source from the signals of other nearby sources of comparable strength. Thus, when the position of one source intercepts the main lobe, its measured signal can be significantly higher than the signals received from the other sources.

Attenborough[7] has discussed the feasibility of designing the impedance of porous ground to control railway noise. Thus, it would also be feasible to use the ground structure described here for the control of low-frequency noise. For example, the control of transformer noise is difficult because of the low frequencies involved (100 or 120 Hz and their harmonics). By constructing a number of strips in the ground around the transformer, tuned to the desired frequencies, the energy is trapped along the strips by the formation of surface waves. The surfaces waves could then be quenched using appropriate resistive material or an active technique. Estimating the amount of noise reduction expected at longer distances from the transformer is beyond the scope of this paper.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Tony Embleton for his suggestions regarding this work. The technical assistance of René St. Denis in constructing the structured ground and collecting data is appreciated.

[1] J. W. Benson, Y.-L. Li, and G. W. Swenson, Jr., ''A baffle-type directional microphone,'' J. Acoust. Soc. Am. **95**, 2536–2538 (1994).
[2] L. M. Brekhovskikh, ''Surface waves in acoustics,'' Sov. Phys. Acoust. **5**, 3–12 (1959).
[3] W. Zhu, G. A. Daigle, and M. R. Stinson, ''Experimental and numerical study of air-coupled surface waves generated above strips of finite impedance,'' J. Acoust. Soc. Am. **114**, 1243–1253 (2003).
[4] G. A. Daigle, M. R. Stinson, and D. I. Havelock, ''Experiments on surface waves over a model impedance plane using acoustical pulses,'' J. Acoust. Soc. Am. **99**, 1993–2005 (1996).
[5] T. F. W. Embleton, J. E. Piercy, and G. A. Daigle, ''Effective flow resistivity of ground surfaces determined by acoustical measurements,'' J. Acoust. Soc. Am. **74**, 1239–1244 (1983).
[6] D. G. Albert, ''Observation of acoustic surface waves in outdoor sound propagation,'' J. Acoust. Soc. Am. **113**, 2495–2500 (2003).
[7] K. Attenborough, ''On the feasibility of railway noise reduction by using local porous surfaces,'' Proc. INTER-NOISE, Leuven, 1993, Vol. 93, pp. 507–511.

An effective fluid model for landmine detection using acoustic to seismic coupling

Doru Velea
Planning Systems Inc., 12030 Sunrise Valley Dr., Reston Plaza I, Suite 400, Reston, Virginia 20191

Roger Waxler[a)]
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

James M. Sabatier
National Center for Physical Acoustics, University of Mississippi, University, Mississippi 38677

(Received 21 June 2003; revised 5 January 2004; accepted 29 January 2004)

A model is developed to describe the response of the ground to airborne sound in the presence of a buried landmine. The model describes both the near-field acoustic wave phenomena as well as the mechanical dynamics of the mine itself. Since buried landmines are typically close to the surface of the soil, the induced ground vibration is the result of near-field scattering so that classical asymptotic scattering theory cannot be used. Instead, the near-field problem is solved in a right circular waveguide with rigid walls. The waveguide contains air in the upper half, soil in the lower half, and a buried mine placed concentrically on the waveguide's axis. The advantage of a waveguide model over infinite space models is that the resulting computer models are much more straightforward to code and run much more quickly. As the radius of the waveguide increases, the results of the waveguide model converge to those of the infinite space model. For low frequencies (a few hundred Hz) this convergence is quite rapid: it is found that a waveguide radius of 10 times the mine radius is more than sufficient. The results obtained are found to qualitatively explain the phenomena observed in the field. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1691038]

PACS numbers: 43.28.En, 43.28.Js, 43.20.Fn [LCS] Pages: 1993–2002

I. INTRODUCTION

Methods for detecting buried landmines using airborne sound have been developed recently.[1–3] Typically, a loudspeaker is used to insonify the ground, and the resulting vibrations of the surface of the ground is measured. A depiction of a typical experiment is found in Fig. 1. It is found that above a buried landmine the amplitude of the ground's vibration is much greater than it is away from the mine. In Fig. 2 some typical results for antitank mines are displayed. The velocity amplitude is measured with a laser Doppler vibrometer (LDV) and is depicted on a color scale plot. The mine is a VS 1.6 antitank mine buried at 3 in. (0.075 m); its radius and height are about 0.07 and 0.1 m, respectively. The broadcast sound is typically broadband, covering a frequency range from around 80 to 300 Hz. The frequency response of the ground vibration at a point directly above the mine is depicted in Fig. 3.

There are several distinctive features in the data. The enhancement of the ground's vibration due to the mine remains localized above the mine itself, indicating that the observed acoustic scattering is near field only, and does not propagate away from the mine to the far field. The frequency spectrum of the ground velocity above the mine is not broadband, but is peaked about certain frequencies (a large peak is visible in Fig. 3 near 100 Hz). Further, there are frequency bands in which the ground's vibration above the mine appears to be suppressed. Such a suppression is visible in Fig. 3 near 180 Hz. In Fig. 4 the data used to produce Fig. 3 are filtered to the frequency band from 180 to 210 Hz. The supression seen in Fig. 3 is seen to be accompanied by a ring of enhanced vibration surrounding the mine.

It has been suggested by Donskoy *et al.*[4,5] that the enhanced vibration is due to a mechanical resonance of the mine which is being excited by the acoustic field. It is the presence of this mechanical resonance which is responsible for the observed peak in the frequency response of the ground vibration. A lumped-element model has been developed[4,5] to describe a one-dimensional approximation to the mine–soil system. Rigorously, this type of model is a low-frequency approximation to the one-dimensional model obtained by considering normally incident sound on an impedance plane whose impedance is chosen to match the mechanical response of the top of the mine (called in this paper the ''infinite mine model;'' see Fig. 12). Such a model does not include any scattering effects but should give qualitative

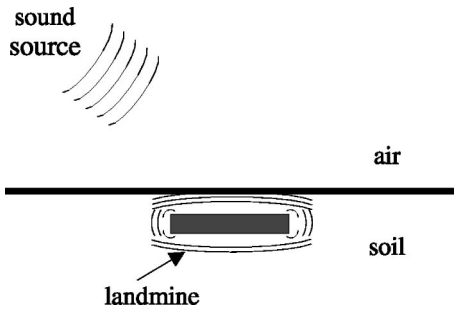


FIG. 1. The typical experimental configuration.

[a)]Electronic mail: rwax@olemiss.edu

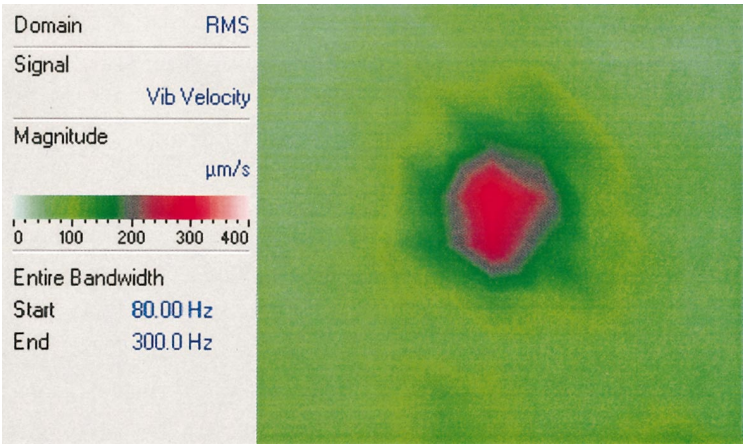


FIG. 2. LDV scan of a VS1.6 antitank mine buried at a depth of 3 in. The magnitude of the soil surface vibration is depicted. The region depicted is a 1-m square.

results about the vibration amplitude directly above the center of the mine as long as higher dimensional effects (in this case radial) can be ignored. In particular, a shift due to mass loading by the soil in the resonant peak produced by the mechanical resonance of the mine is predicted: the resonance frequency is predicted to decrease monotonically with depth.

Donskoy's lumped-element model has been studied in more detail by Yu *et al.*,[6] who pointed out that, in addition to the enhanced vibration in a frequency band containing the mechanical resonance frequency, there is a frequency at which the mechanical vibration of the mine is suppressed (in the lumped-element model there is a frequency at which the vibration amplitude is zero) due to destructive interference between the induced elastic field in the soil and the mechanical vibration of the mine. They then suggested that it is this effect which is responsible for the observed rings of the sort depicted in Fig. 4.

In this paper a simple scattering model is developed. In order to obtain a tractable mathematical problem the sound field is assumed to be incident normally on the soil, and the soil is modeled as an effective fluid with a single compression mode. Although this is an overly simplistic model, it will be seen that a qualitative explanation of the basic phenomena emerges. The method used to solve the resulting mathematical problem, discussed in Sec. III, depends critically on the symmetry preserved by assuming normally incident sound as well as on the spherical spreading of the sound field in the effective fluid.

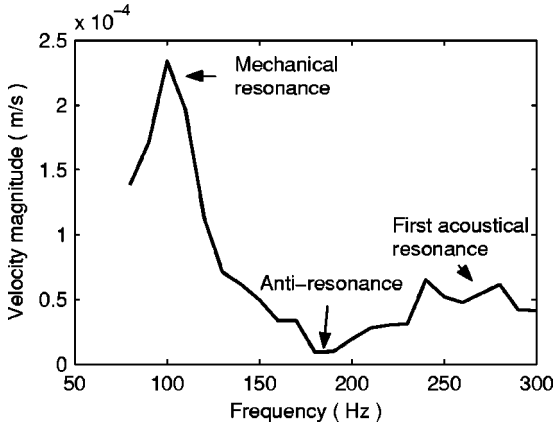


FIG. 3. Frequency response of the point on the soil surface directly above the center of the mine for the velocity whose magnitude is depicted in Fig. 2.

Soils are more often modeled as poroelastic solids which support vector shear and fast and slow compression modes.[7] The fast wave corresponds to the compressional elastic wave that propagates in elastic solids, while the slow wave corresponds to the diffusive wave that propagates in rigid porous media. It is unclear at present what the influence of the shear is and what the relative importance of the fast and slow waves is. Thus, it is unclear how good an approximation an effective fluid model is. Further, it is also unclear what dispersion relation should be used to characterize the effective fluid.

To attain the greatest possible generality, no particular dispersion relation is assumed. Instead, we will simply assume that, given a particular frequency, there is a corresponding wave number. Thus, we consider this model to be a generalization of the rigid-porous model used by Berry *et al.*[8] In the concrete examples presented it is assumed that the fluid is dispersionless. It would be straightforward, however, to include any desired relation between wave number and frequency. It is also straightforward to include attenuation in the model by specifying the imaginary part of the wave number. Wave numbers with negligible imaginary parts are typical of the fast wave, while wave numbers with comparable real and imaginary parts are typical of the slow wave.

It is found that the observed scattered field is short range because it is significant only in the near field. Introducing attenuation in the soil into the model plays little role in the range of the scattered field on the air/soil interface, but does decrease the amplitude somewhat. Further, it is found that the resonance frequency corresponding to the mechanical resonance of the mine decreases with depth as predicted by the one-dimensional approximation; however, it reaches a minimum, ceasing to change after a certain depth is reached.

At somewhat higher frequencies than the mechanical resonance, one finds another resonance. This is largely due to resonant behavior in the layer of soil between the top of the mine and the surface of the soil and is essentially the quarter-wavelength resonance, perturbed by the impedance of the top of the mine, that one would find between a rigid surface and the surface of the soil. This resonance will be referred to in

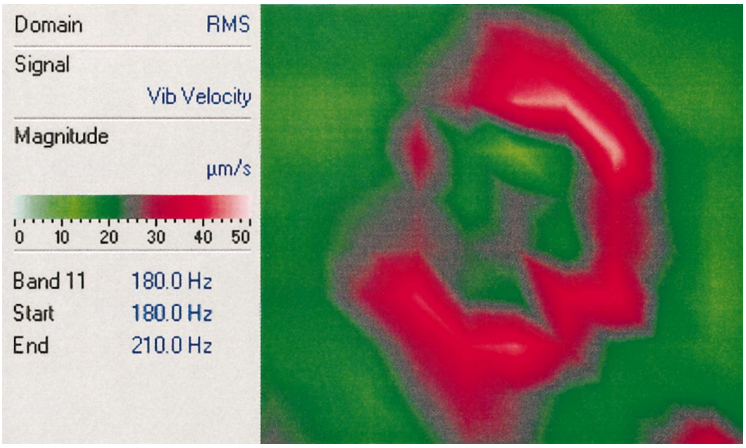


FIG. 4. Visualization of the ring pattern observed in the narrow-band response. The magnitude of the soil surface vibration from Fig. 2 is restricted to the frequency band from 180 to 210 Hz. The region depicted is a 0.9-m square.

this paper as the first acoustic resonance. The suppression of the mechanical vibration of the mine, discovered by Yu *et al.*,[6] is shown to be an antiresonance which occurs between the fundamental mechanical resonance and the first acoustic resonance. For frequencies near this antiresonance frequency the scattered field has the form of a ring of higher velocity surrounding the position of the mine, much as has been seen in the data.

Finally, on introducing attenuation into the soil model one finds that the position of both the mechanical resonance and the first acoustic resonance seem to vary little, although the magnitude of the resonant peaks decrease. The mechanical resonance persists, albeit with smaller amplitude. The first acoustic resonance, however, rapidly broadens and vanishes with increasing attenuation. The depth of the antiresonance minimum decreases; however, the ring-shaped scattering pattern persists.

II. DESCRIPTION OF THE MODEL

In this section the model discussed above is introduced. The geometry of the model is depicted in Fig. 5. Let z measure vertical displacement. For $z>0$ one has air with (real-valued) mean density ρ_0 and sound speed c_0. For $z<0$ one has an effective fluid (a simplified soil) with (possibly complex-valued) mean density ρ_1 and sound speed c_1. In the calculations presented here the values $c_0=340\,\mathrm{m/s}$, $\rho_0=1.2\,\mathrm{kg/m^3}$ appropriate to air, and the nominal values $c_1=160\,\mathrm{m/s}$ and $\rho_1=1400\,\mathrm{kg/m^3}$ for the soil are used.

In this paper the mine is assumed to be a compliant object with one mechanical degree of freedom. To simplify further it is modeled as a right circular cylinder with radius a and height h, buried to a depth d. The mechanical degree of freedom is modeled as a piston on the top of the mine. The piston can oscillate harmonically in the vertical direction with effective dynamical mass M, spring constant S, and resistance R. Let ξ represent the displacement from equilibrium of the piston. In the calculations presented here the values $S=2.5\cdot 10^7\,\mathrm{N/m^3}$, $M=12\,\mathrm{kg/m^2}$, and $R=1700\,\mathrm{N\cdot s/m^3}$ are used. These are the parameters for which the piston has the acoustic impedance of the fundamental mode of a VS1.6 antitank mine.[5] Note that with such parameters an unburied mine has a mechanical resonance at 230 Hz and a quality factor of $Q=10.2$.

Assuming a time-harmonic pressure field (in cylindrical coordinates) $P(r,\theta,z)e^{-i\omega t}$, the amplitude P satisfies the Helmholtz equation

$$(\nabla^2+k_j^2)P=0, \tag{1}$$

with $k_j=\omega/c_j$, $j=0$ for $z>0$ and $j=1$ for $z<0$. At $z=0$ one has continuity of pressure

$$P|_{z=0^+}=P|_{z=0^-}, \tag{2}$$

and continuity of particle velocity, which leads to

$$\frac{1}{\rho_0}\frac{\partial P}{\partial z}\bigg|_{z=0^+}=\frac{1}{\rho_1}\frac{\partial P}{\partial z}\bigg|_{z=0^-}. \tag{3}$$

The wave number in the air, k_0, is real valued, but the wave number in the soil model, k_1, in general is not. It will often be convenient to write k_1 as a sum of its real and imaginary parts

$$k_1=q+i\alpha.$$

The sides and bottom of the mine are assumed to be rigid, so that

$$\frac{\partial P}{\partial r}\bigg|_{r=a}=0, \tag{4}$$

and, for $r\leqslant a$

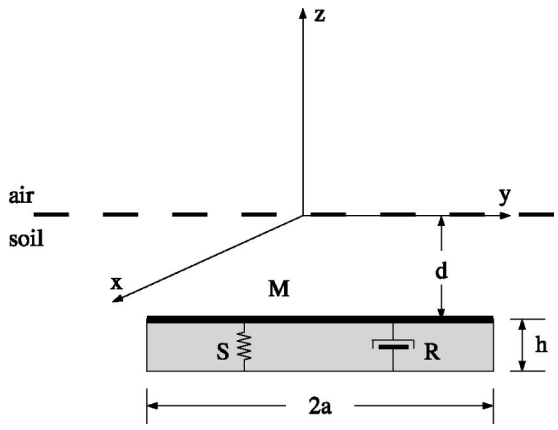


FIG. 5. The geometry of the model.

$$\left.\frac{\partial P}{\partial z}\right|_{z=-d-h}=0. \tag{5}$$

On the top of the mine the velocity of the piston, $\partial\xi/\partial t$, must equal the particle velocity in the soil so that, for $r\leqslant a$

$$-i\omega\xi=\frac{1}{i\omega\rho_1}\left.\frac{\partial P}{\partial z}\right|_{z=-d}.$$

The acoustic field in the soil drives the piston so that at the top of the mine one has Newton's law

$$\int_0^a\int_0^{2\pi}P(r,\theta,-d)r\,d\theta\,dr=(-\omega^2M-i\omega R+S)\xi.$$

It follows from the previous two equations that

$$\int_0^a\int_0^{2\pi}P(r,\theta,-d)r\,d\theta\,dr=-\Omega(\omega)\left.\frac{\partial P}{\partial z}\right|_{z=-d}, \tag{6}$$

where

$$\Omega(\omega)=\frac{\omega^2M+i\omega R-S}{\omega^2\rho_1}. \tag{7}$$

Note that $-i\omega\rho_1\Omega$ is the acoustic impedance of the piston.

The pressure field is divided into unperturbed and scattered parts, $P=P_0+P_S$. The unperturbed part is planar. Above the interface it contains a downward-going part (the source) and an upward-going reflected part. Below the interface it contains a downward-going transmitted part. One has

$$P_0(z)=A\begin{cases}e^{-ik_0z}+\dfrac{\rho_1c_1-\rho_0c_0}{\rho_0c_0+\rho_1c_1}e^{ik_0z} & \text{if } z>0,\\ \dfrac{2\rho_1c_1}{\rho_0c_0+\rho_1c_1}e^{-ik_1z} & \text{if } z<0\end{cases},$$

where A is the amplitude of the incoming field.

The scattered part is the solution of the Helmholtz equations (1) which satisfies the boundary conditions (2) and (3) at the air/soil interface $z=0$, as well as

$$\left.\frac{\partial P_s}{\partial r}\right|_{r=a}=0, \tag{8}$$

for $-d-h<z<-d$

$$\left.\frac{\partial P_s}{\partial z}\right|_{z=-d-h}=-P_0'(-d-h), \tag{9}$$

and

$$2\pi\int_0^a P_s(r,-d)r\,dr+\Omega(\omega)\left.\frac{\partial P_s}{\partial z}\right|_{z=-d}=-\pi a^2P_0(-d)-\Omega(\omega)P_0'(-d), \tag{10}$$

for $0\leqslant r\leqslant a$. For $a<r$ both P_s and $\partial P_s/\partial z$ are continuous. In (10) we have used the fact that P_s will be independent of θ.

Of principal interest here is the velocity amplitude of the soil surface, and in particular the enhancement due to the presence of the mine. The velocity amplitude $v_1(r,0)$ of the soil surface is given by

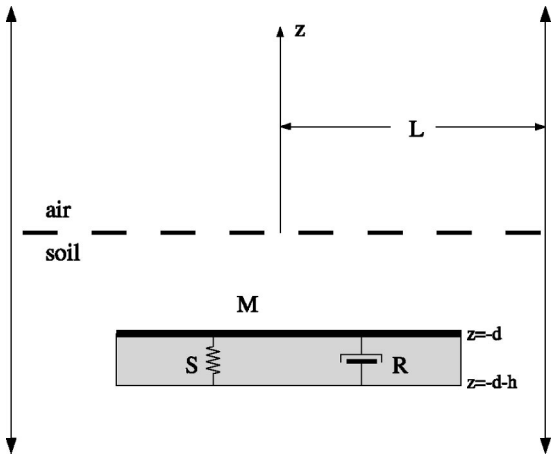


FIG. 6. The computational domain.

$$v_1(r,0)=\frac{1}{i\omega\rho_1}\left.\frac{\partial P}{\partial z}\right|_{z=0}=\frac{1}{i\omega\rho_1}\left(P_0'(0)+\left.\frac{\partial P_s}{\partial z}\right|_{z=0}\right).$$

A measure of the influence of the mine is the ratio, $U(r)$, of the soil surface velocity in the presence of the mine to the velocity in the absence of the mine. One finds

$$U(r)=1+\frac{1}{P_0'(0)}\left.\frac{\partial P_s}{\partial z}\right|_{z=0}. \tag{11}$$

III. THE METHOD OF SOLUTION

We need to solve for the scattered field P_s defined in the previous section. Since we are concerned with the near field, the usual far-field asymptotic forms of scattering theory cannot be used. Instead, we will resort to a brute force numerical scheme. A simple and efficient scheme can be devised by replacing the full-space domain by a circular waveguide. The z axis is taken to be the symmetry axis of the waveguide and is centered at the center of the mine. The radius of the waveguide is taken to be L. The walls of the waveguide are taken to be rigid so that a plane wave can propagate in it. The geometry is depicted in Fig. 6. In the limit $L\to\infty$ one recovers the full-space results. The usefulness of such an approach depends on how large L needs to be and, given L, on how straightforward the resulting waveguide problem is.

For finite L one has a waveguide with discontinuities at the top and bottom of the mine: the waveguide changes abruptly from circular to annular. Our approach will be to expand the scattered field in the transverse eigenfunctions of a circular waveguide for $z>-d$ and $z<-d-h$ and in the transverse eigenfunctions of an annular waveguide for $-d-h<z<-d$. The full field is then determined by mode matching at $z=-d$ and $z=-d-h$.

More explicitly (the necessary details about the eigenvalues and eigenfunctions are given in the Appendix) if ψ_j, λ_j are the transverse eigenfunctions and eigenvalues, respectively, for a circular waveguide and ϕ_j, ϵ_j are the transverse eigenfunctions and eigenvalues, respectively, for a rigid-walled annular waveguide, then

$$P_s(r,z)=\begin{cases}\sum_{j=0}^{\infty} A_j(z)\psi_j(r) & \text{if } z>-d,\\ \sum_{j=0}^{\infty} B_j(z)\phi_j(r) & \text{if } -d-h<z<-d,\\ \sum_{j=0}^{\infty} C_j(z)\psi_j(r) & \text{if } z<-d-h,\end{cases} \tag{12}$$

where

$$A_j(z)=a^{(j)}\begin{cases} e^{i\sqrt{k_0^2-\lambda_j}z} & \text{for } z>0,\\ \frac{1}{2}\left(1+\frac{\rho_1\sqrt{k_0^2-\lambda_j}}{\rho_0\sqrt{k_1^2-\lambda_j}}\right)e^{i\sqrt{k_1^2-\lambda_j}z}+\frac{1}{2}\left(1-\frac{\rho_1\sqrt{k_0^2-\lambda_j}}{\rho_0\sqrt{k_1^2-\lambda_j}}\right)e^{-i\sqrt{k_1^2-\lambda_j}z} & \text{for } -d<z<0,\end{cases}$$

$$B_j(z)=b_1^{(j)}e^{i\sqrt{k_1^2-\epsilon_j}z}+b_2^{(j)}e^{-i\sqrt{k_1^2-\epsilon_j}z},$$

and

$$C_j(z)=c^{(j)}e^{-i\sqrt{k_1^2-\lambda_j}z}.$$

The coefficients $a^{(j)}$, $b_1^{(j)}$, $b_2^{(j)}$, and $c^{(j)}$ are the unknowns in this formulation and are determined by the boundary conditions (9) and (10) and the conditions that P_s and $\partial P_s/\partial z$ are continuous at $z=-d$ and $z=-d-h$ for $a<r<L$. Note that (see the Appendix) the eigenvalues λ_j and ϵ_j increase without bound as j increases, having asymptotic forms given by

$$\lambda_j \approx \frac{\pi^2\left(j+\frac{1}{4}\right)^2}{L}, \tag{13}$$

and

$$\epsilon_j \approx \frac{\pi^2 j^2}{L-a}, \tag{14}$$

for large j.

For sufficiently large j the modes become evanescent in the sense that their coefficients behave exponentially as a function of z and do not propagate to large distances from the soil surface. However, since the soil surface is in the near field, it is the contribution from the evanescent modes that determines the velocity at the soil surface.

At the top of the mine, at $z=-d$, one has from the continuity of the pressure field with $a<r<L$

$$\sum_{j=0}^{\infty} A_j(-d)\psi_j(r)=\sum_{j=0}^{\infty} B_j(-d)\phi_j(r), \tag{15}$$

and from the continuity of the velocity field

$$\sum_{j=0}^{\infty} A_j'(-d)\psi_j(r) = \begin{cases} -\frac{2\pi}{\Omega(\omega)}\left(\frac{a^2}{2}P_0(-d)+\sum_{j=0}^{\infty} A_j(-d)\chi_j\right) & \\ \quad -P_0'(-d) & \text{if } 0<r<a,\\ \sum_{j=0}^{\infty} B_j'(-d)\phi_j(r) & \text{if } a<r<L,\end{cases} \tag{16}$$

where $\chi_j=\int_0^a \psi_j(r)r\,dr$. At the bottom of the mine, at $z=-d-h$, one has from the continuity of the pressure field with $a<r<L$

$$\sum_{j=0}^{\infty} C_j(-d-h)\psi_j(r)=\sum_{j=0}^{\infty} B_j(-d-h)-h\phi_j(r), \tag{17}$$

and from the continuity of the velocity field

$$\sum_{j=0}^{\infty} C_j'(-d-h)\psi_j(r) = \begin{cases} -P_0'(-d-h) & \text{if } 0<r<a,\\ \sum_{j=0}^{\infty} B_j'(-d-h)\phi_j(r) & \text{if } a<r<L.\end{cases} \tag{18}$$

Multiplying the two pressure equations, Eqs. (15) and (17), by $\phi_k(r)r$, integrating from $r=a$ to $r=L$, multiplying the two velocity equations, Eqs. (16) and (18), by $\psi_k(r)r$, and integrating from $r=0$ to $r=L$, one obtains four sets of equations. Defining the overlap matrix

$$T_{kj}=\int_a^L \phi_k(r)\psi_j(r)r\,dr,$$

one has (here, δ_{kj} is the Kronecker delta)

$$\sum_{j=0}^{\infty}\left(T_{kj}A_j(-d)-\delta_{kj}B_j(-d)\right)=0, \tag{19}$$

$$\sum_{j=0}^{\infty}\left(\frac{2\pi}{\Omega(\omega)}\chi_k\chi_j A_j(-d)+\delta_{kj}A_j'(-d)-T_{jk}B_j'(-d)\right)$$
$$=-\left(\frac{\pi a^2}{\Omega(\omega)}P_0(-d)+P_0'(-d)\right)\chi_k, \tag{20}$$

$$\sum_{j=0}^{\infty}\left(T_{kj}C_j(-d-h)-\delta_{kj}B_j(-d-h)\right)=0, \tag{21}$$

$$\sum_{j=0}^{\infty}\left(T_{jk}B_j'(-d-h)-\delta_{kj}C_j'(-d-h)\right)$$
$$=P_0'(-d-h)\chi_j. \tag{22}$$

Truncating the infinite sums in Eqs. (19) through (22) at some finite index $j=N-1$, one obtains $4N$ linear equations for the $4N$ unknown coefficients $a^{(j)}$, $b_1^{(j)}$, $b_2^{(j)}$, $c^{(j)}$ with $j\in\{0,1,2,...,N-1\}$. One could choose a different truncation index in each of the regions; above the mine $z>-d$, around the mine $-d-h<z<-d$, and below the mine $z<-d-h$; however, we have not investigated this possibility. These can be easily solved numerically and then the resulting coefficients can, in principle, be substituted into (12), giving an approximate solution for the pressure field.

This approximation cannot be expected to be uniformly valid for all values of z. In particular, at the top and bottom of the mine we expect the convergence with increasing N to be poor because of the abrupt transition from a circular to an annular duct. Further, a sufficient number of the evanescent modes must be kept to correctly account for the interaction between the surface of the soil and the top of the mine. However, we are interested in the velocity field at the surface of the soil, and not on the mine itself. Since the distance from the mine to the soil surface, d, is greater than zero, convergence can be achieved, but the approach will fail as $d\to 0$. Similarly, in the annular region surrounding the mine, enough of the evanescent modes must be kept to account for any interaction between the top and bottom of the mine.

For large j the z dependence of the jth mode in the region between the mine and the soil surface is given by $e^{\pm\sqrt{\lambda_j-k_1^2}z}$. Thus, given a burial depth d, N must be large enough so that $e^{-\sqrt{\lambda_j-k_1^2}d}$ is sufficiently small. If

$$e^{-\sqrt{\lambda_N-k_1^2}d}<\eta$$

for some small parameter η (in practice $\eta=0.0001$ is found to be more than sufficient), then the asymptotic form (13) of the eigenvalues λ_j shows that it is sufficient to choose

$$N\geqslant\frac{L}{\pi}\sqrt{k_1^2+\left(\frac{\ln(1/\eta)}{d}\right)^2}.$$

Similarly, in the region surrounding the mine, for large j the z dependence of the jth mode is given by $e^{\pm\sqrt{\epsilon_j-k_1^2}z}$. The asymptotic form (14) of the eigenvalues ϵ_j shows that it is sufficient to choose

$$N\geqslant\frac{L-a}{\pi}\sqrt{k_1^2+\left(\frac{\ln(1/\eta)}{h}\right)^2}.$$

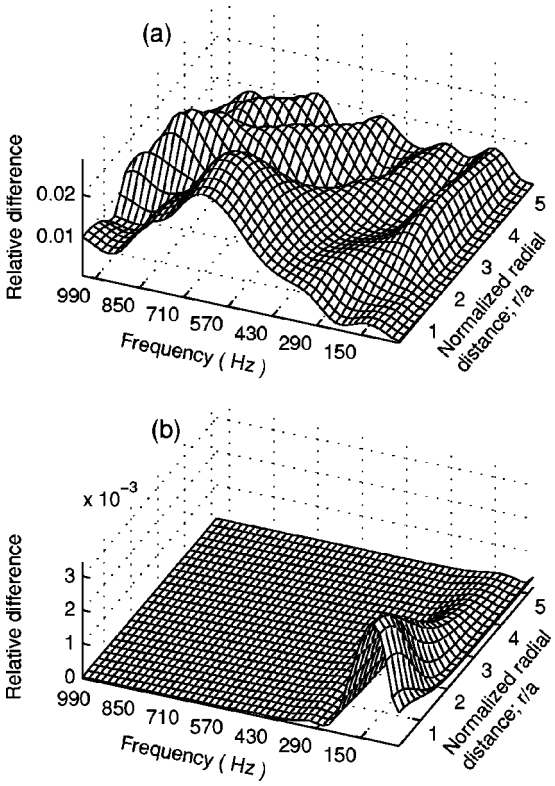


FIG. 7. Relative error, as given by Eq. (23), for the predicted ground velocity (a) with no attenuation and (b) with $\mathrm{Re}\,k=\mathrm{Im}\,k$. The depth is 10 cm.

In practice, N is chosen large enough to satisfy both of these criteria.

Finally, L must be chosen large enough to recover the infinite space solution for the velocity on the soil surface. In addition, there are values of L that should be avoided because they support radial resonances in the duct. These values are determined by the condition $k_1=\sqrt{\lambda_j}$ for $z<0$. For $z>0$ such radial resonances occur as well but have little effect because the scattered field is small above the soil. In practice, we avoid any L close to the critical values $z_{1j}/k_1\approx\pi j/k_1$. As the frequency increases larger values of L/a are needed to converge to the infinite space result. We find that choosing $L/a=15$ is sufficient up to the kHz range.

As a specific example the relative velocity $U(r)$ is calculated for $d=10$ cm. In these runs L/a is kept fixed and a set of closely spaced frequencies is chosen ranging from 10 Hz up to 990 Hz. The frequencies sampled are not evenly spaced. Rather, they are chosen to avoid having any critical values for L.

The convergence to the infinite space results is demonstated in Figs. 7(a) and 7(b). In these figures the relative differences between the predicted values obtained for the relative velocity $U(r)$ [recall Eq. (11)] with $L/a=15$ and $L/a=100$,

$$\left|\frac{U(r)|_{L/a=100}-U(r)|_{L/a=15}}{U(r)|_{L/a=100}}\right|, \tag{23}$$

are plotted as functions of frequency and radial distance from the center of the mine. In Fig. 7(a) k_1 is taken to be real, so that there is no attenuation in the soil, while in Fig. 7(b) the real and imaginary parts of k_1 are taken to be equal to each

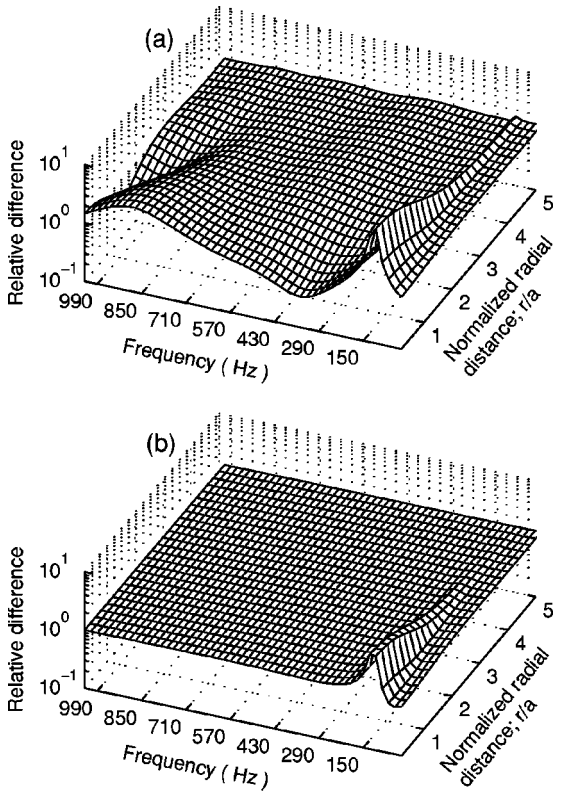


FIG. 8. The normalized ground velocity magnitude, in dB relative to 1.0. (a) with no attenuation and (b) with $\mathrm{Re}\,k=\mathrm{Im}\,k$. The depth is 10 cm.

other, $\mathrm{Re}\,k_1=\mathrm{Im}\,k_1$, so that there is significant attenuation in the soil. In Figs. 8(a) and 8(b) the magnitude of the relative velocity $U(r)$ (with $L/a=100$) is plotted, again as functions of frequency and radial distance from the center of the mine, and again with $\mathrm{Re}\,k_1=0$ and $\mathrm{Re}\,k_1=\mathrm{Im}\,k_1$, respectively.

Note that with no attenuation the relative error is no greater than 2% [Fig. 7(a)], and that in a frequency band, 500 to 700 Hz, where the relative velocity itself is small [Fig. 8(a)]. With attenuation the convergence is even better. Thus, we may conclude that $L/a=15$ is sufficient up to about 1 kHz. It should be pointed out that run times are quite reasonable. On a desktop personal computer the computation time needed for a single frequency is on the order of a few seconds.

IV. RESULTS

The foremost predictions of this model are summarized in Figs. 8 through 13. In Fig. 8 the normalized surface velocity magnitude $|U(r)|$ is plotted as a function of frequency and distance from the center of the mine. In Fig. 8(a) no attenuation is included, while in Fig. 8(b) there is 100% attenuation, $\mathrm{Re}\,k_1=\mathrm{Im}\,k_1$. (In general, in the figures that follow the amount of attenuation in the model is noted by indicating how large the imaginary part of the wave number is compared to the real part. Thus, the label x% attenuation indicates that $\mathrm{Im}\,k_1=0.01\cdot x\cdot\mathrm{Re}\,k_1$.) In Fig. 9 the magnitude and phase of the normalized surface velocity directly above the center of the mine is plotted as a function of frequency. No attenuation has been introduced in this plot. The features

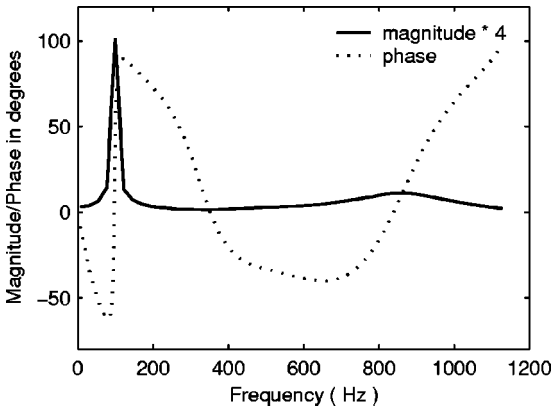


FIG. 9. The normalized ground velocity magnitude and phase above the center of the mine. The depth is 10 cm.

visible in the plots in Fig. 8 and Fig. 9 correspond qualitatively to what is found experimentally in the field as in Fig. 2, Fig. 3, and Fig. 4.

The mechanical resonance is visible at about 100 Hz. It is this resonance that is responsible for the enhancements of the surface velocity seen in the field (recall Fig. 2 and Fig. 3). The mechanical resonance is studied further in Fig. 10, in which the normalized surface velocity is plotted as a function of radial distance from the center of the mine at the resonance frequency of the mine–soil system. The response of a compliant mine is compared to that of a rigid mine. The solid curves represent the response in soil without attenuation, the dashed curves in soil with 100% attenuation, $\mathrm{Re}\,k_1=\mathrm{Im}\,k_1$. Note that while the attenuation decreases the magnitude of the peak, the overall form of the spatial response in largely unaffected. Further, Fig. 8 shows that this peak doesn't broaden much.

There is a second peak in Fig. 8(a) and Fig. 9 at about 860 Hz. This peak is the first acoustic resonance, arising from resonant behavior in the soil between the top of the mine and the surface of the soil. In Fig. 11 the normalized surface velocity is plotted again as a function of radial distance from the center of the mine but now at the resonance frequency of the first acoustic resonance. Note that with increased attenuation this peak vanishes.

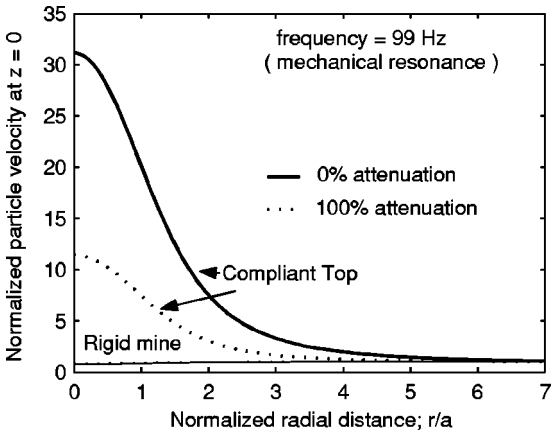


FIG. 10. Radial profile of the effect of the mechanical resonance on the ground velocity with no attenuation and with $\mathrm{Re}\,k=\mathrm{Im}\,k$. The depth is 10 cm.

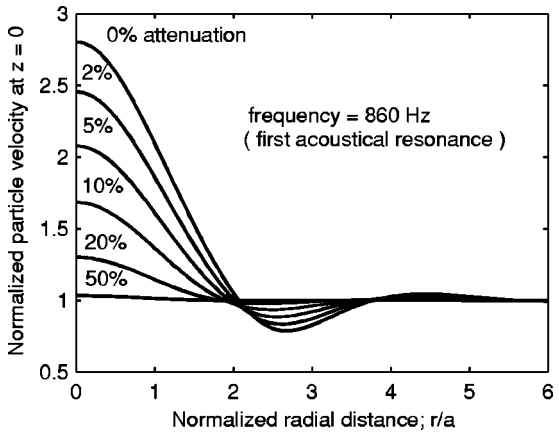


FIG. 11. Radial profile of the first acoustic resonance and the effect of increasing attenuation. The depth is 10 cm.

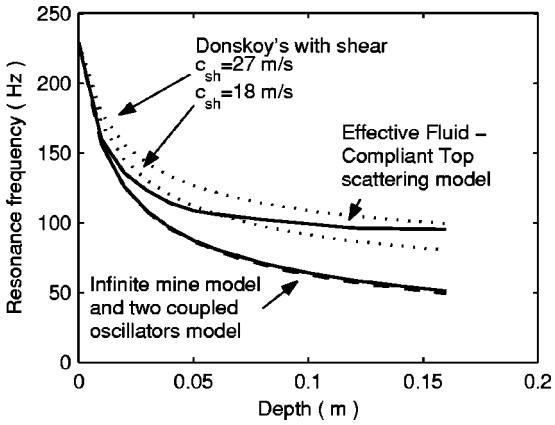


FIG. 13. The shift in the mechanical resonance frequency with increasing depth as predicted by the effective fluid model is compared to the predictions of Donskoy's lumped-element model.

Between the mechanical and first acoustic resonance is the antiresonance predicted by Yu *et al.*,[6] found in data such as in Fig. 4. This antiresonance appears in Fig. 8(a) as a broad valley between 290 and 430 Hz and in Fig. 9 as a zero of the phase. In Fig. 12 the normalized surface velocity as a function of radial distance is plotted at the antiresonance frequency. Note that attenuation reduces the magnitude of the effect, but it does not remove it completely.

Finally, the frequency at which the mechanical resonance is seen depends on the depth to which the mine is buried. In Fig. 13 the prediction from the effective fluid scattering model for this dependence is compared to those from various versions, with and without shear in the soil, of Donskoy's lumped-element model.[5] In addition, the response from the infinite mine model (a radially infinite mine) in an effective fluid is plotted.

Consider first the dependence of the fundamental mechanical resonance frequency on depth in the lumped-element model without shear (this would be the lumped-element approximation to the effective fluid model). The prediction of this lumped-element model is almost identical to that of the infinite mine model. These one-dimensional models predict a resonance frequency which decreases steadily with increasing depth. Putting shear into the lumped-element model has the effect of increasing this frequency, but still predicts that the resonance frequency will continue to decrease as the depth increases. The physical interpretation which arises is that the resonance frequency is lowered due to mass loading of the mine by the soil column above it.

The effective fluid scattering model, however, predicts a fundamental mechanical resonance frequency which decreases with depth, but which eventually reaches a minimum, here at about 12 cm, remaining more or less constant as the depth increases further. One presumes that, with increasing depth, the radial direction becomes more and more important, moderating the mass loading due to the soil column above the mine until it reaches an effective constant at some depth. Since the lumped-element models are intrinsically one-dimensional, they don't account for scattering in the radial direction. In Fig. 13 it is shown that the predictions for the fundamental mechanical resonance frequency from the lumped-element model without shear and from the effective fluid model agree with each other asymptotically as the depth approaches zero; however, they differ significantly for nonzero depth. In general, one may expect the lumped-element models to be valid only for very shallow mines for which soil motion in the radial direction is unimportant.

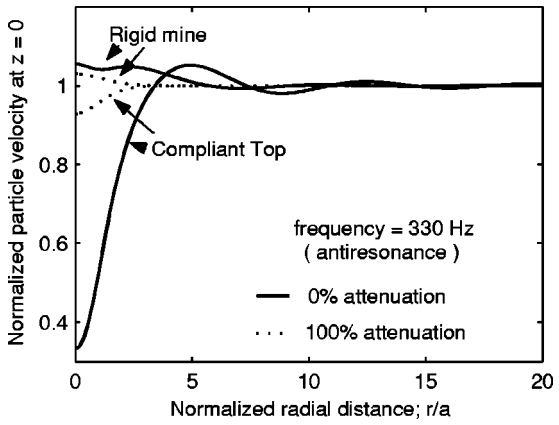


FIG. 12. Radial profile of the antiresonance with no attenuation and with Re k=Im k. The depth is 10 cm.

V. CONCLUSIONS

The method developed here gives an easily coded and efficient algorithm to describe the near field scattering of sound by a buried landmine. Modeling the soil with an effective fluid ignores the effects of shear as well as porosity, and modeling a landmine as a rigid cylinder with a piston on top oversimplifies the mechanical response of an actual landmine. Nonetheless, the model gives qualitative explanations of the observed phenomena and makes some definite predictions.

First it is verified that, as predicted by Donskoy, the enhancement of the ground's vibration over a mine is due to the excitation of a mechanical resonance of the mine itself. In addition, the observed short range of the scattering is shown to be a feature of the near field and the inefficiency of the acoustic to seismic coupling. Further, attenuation is shown to have little effect on the range of the scattered field.

Next it is verified, as conjectured by Yu *et al.*,[6] that there is a frequency at which the vibration of the mine is at a minimum. It is shown that this minimum is an antiresonance between the resonance produced by the mechanical vibration of the mine and an acoustic resonance in the soil layer above the mine. Near this antiresonance frequency the ground velocity magnitude has the form of a ring.

Finally, the effect of varying the depth to which the mine is buried is investigated. It is shown that scattering in the radial direction becomes important with increasing mine depth. For shallow mines the frequency of the mechanical resonance shifts due to mass loading by the soil. As the depth of the mine increases, the mechanical resonance frequency eventually reaches a minimum and ceases to shift. Further, the magnitude of the resonance peaks and the depth of the antiresonance minimum decrease with increasing depth.

APPENDIX: EIGENFUNCTIONS AND EIGENVALUES

The set of eigenfunctions $\{\psi_j(r)\}$, $j=0,1,2,...$, for the cylindrically symmetric part of the acoustic field in a rigid circular waveguide of radius L satisfies[9]

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\lambda_j\right)\psi_j(r)=0, \quad \text{for } 0\leqslant r\leqslant L,$$

with the boundary condition

$$\left.\frac{\partial\psi_j(r)}{\partial r}\right|_{r=L}=0.$$

The functions $\psi_j(r)$ have the well-known form

$$\psi_j(r)=N_j J_0(\sqrt{\lambda_j}r),$$

where J_0 is the zeroth-order Bessel function of the first kind[10] and

$$N_j=\frac{\sqrt{2}}{R|J_0(\sqrt{\lambda_j}L)|}.$$

N_j is the normalization factor determined from the normalization condition

$$\int_0^L \psi_j(r)\psi_m(r)r\,dr=\delta_{jm}, \quad j,m=0,1,2,...,$$

where δ_{jm} is the Kronecker symbol. The eigenvalues λ_j are determined by the boundary condition at $r=L$, which leads to the familiar eigenvalue condition[9]

$$\left.\frac{\partial J_0(\sqrt{\lambda_j}r)}{\partial r}\right|_{r=L}=0.$$

From the identity[10] $J_0'(x)=-J_1(x)$, it follows that

$$\sqrt{\lambda_j}=\frac{z_{1j}}{L},$$

where z_{1j}, $j=0,1,2,...$, denote the zeros of the Bessel function J_1. For $j\gg 1$, z_{1j} has the asymptotic form[10]

$$z_{1j}\to\pi\left(j+\frac{1}{4}\right)-\frac{3}{8\pi\left(j+\frac{1}{4}\right)}+\mathcal{O}\left(\frac{1}{j^2}\right),$$

giving (13).

A similar analysis applies to the annular section of the waveguide. The radial coordinate is restricted to range from a, the mine radius, to L. The corresponding eigenfunctions satisfy

$$\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\epsilon_j\right)\phi_j(r)=0, \quad \text{for } a\leqslant r\leqslant L.$$

Assuming that the side of the mine as well as the waveguide's wall are rigid yields

$$\left.\frac{\partial\phi_j}{\partial r}\right|_{r=a}=\left.\frac{\partial\phi_j}{\partial r}\right|_{r=L}=0 \quad j=0,1,2,....$$

These boundary conditions ensure that the eigenvalues are real, nondegenerate, and form a discrete ordered set $\{\epsilon_0,\epsilon_1,\epsilon_2,...\}$ with the minimum value $\epsilon_0=0$.

It follows from the boundary conditions that the eigenvalues ϵ_j are the zeros of the function

$$f(\epsilon)=J_1(\sqrt{\epsilon}L)Y_1(\sqrt{\epsilon}a)-J_1(\sqrt{\epsilon}a)Y_1(\sqrt{\epsilon}L).$$

The large j asymptotic form for $\{\epsilon_j\}$ can be determined from the asymptotic expansions of the Bessel functions.[10] Noting that, for large x

$$J_1(x)\approx\sqrt{\frac{2}{\pi x}}\cos\left(x-\frac{3\pi}{4}\right)+O(x^{-1}),$$

and

$$Y_1(x)\approx\sqrt{\frac{2}{\pi x}}\sin\left(x-\frac{3\pi}{4}\right)+O(x^{-1}),$$

one finds that, for large ϵ

$$f(\epsilon)\approx\sin(\sqrt{\epsilon}(L-a))+O\left(\frac{1}{\epsilon a}\right).$$

Noting that large ϵ means $j\gg 1$, one obtains (14).

Denoting

$$\mathcal{L}_j(r)=J_0(\sqrt{\epsilon_j}r)-\frac{J_1(\sqrt{\epsilon_j}L)}{Y_1(\sqrt{\epsilon_j}L)}Y_0(\sqrt{\epsilon_j}r),$$

where Y_0 and Y_1 are the zeroth and first-order Bessel function of the second kind,[10] the eigenfunctions are given by

$$\phi_j(r)=\begin{cases}\sqrt{\dfrac{2}{L^2-a^2}} & \text{if } j=0\\ M_j\mathcal{L}_j(r), & j=1,2,3,...\end{cases},$$

with M_j computed from the normalization integral

$$\int_a^L \phi_j(r)\phi_m(r)r\,dr=\delta_{jm}, \quad j,m=0,1,2,...,$$

yielding

$$M_j^2 = \frac{2}{L^2 \mathcal{L}_j^2(L) - a^2 \mathcal{L}_j^2(a)}.$$

[1] J. M. Sabatier and N. Xiang, ''Laser-Doppler based acoustic-to-seismic detection of buried mines,'' in Detection and Remediation Technologies for Mines and Minelike Targets IV, edited by A. C. Dubey, J. F. Harvey, J. T. Broach, and R. E. Dugan (SPIE Proceedings Vol. 3710, 1999), pp. 215–222.

[2] J. M. Sabatier and N. Xiang, ''An investigation of acoustic-to-seismic coupling to detect buried anti-tank landmines,'' IEEE Trans. Geosci. Remote Sens. **39**, 1146–1154 (2001).

[3] N. Xiang and J. M. Sabatier, ''An experimental study on antipersonnel landmine detection using acoustic-to-seismic coupling,'' J. Acoust. Soc. Am. **113**, 1333–1341 (2003).

[4] D. M. Donskoy, N. Sedunov, A. Ekimov, and M. Tsionskiy, ''Optimization of seismo-acoustic land mine detection using dynamic impedances of mines and soil,'' in Detection and Remediation Technologies for Mines and Minelike Targets VI, edited by A. C. Dubey, J. F. Harvey, J. T. Broach, and V. George (SPIE Proceedings Vol. 4394, 2001), pp. 553–562.

[5] D. Donskoy, A. Ekimov, N. Sedunov, and M. Tsionskiy, ''Nonlinear seismo-acoustic land mine detection and discrimination,'' J. Acoust. Soc. Am. **111**, 2705–2714 (2002).

[6] S. Yu, A. Gandhe, T. R. Witten, and R. K. Mehra, ''Physically based method for automatic mine detection using acoustic data,'' in Proc. SPIE Conference on Detection and Remediation Technologies for Mines and Minelike Targets VII, edited by A. C. Dubey, pp. 701–708 (2002).

[7] M. A. Biot, ''Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range,'' J. Acoust. Soc. Am. **28**, 179–191 (1956).

[8] D. L. Berry, S. N. Chandler-Wilde, and K. Attenborough, ''Acoustic scattering by a near-surface obstacle in a rigid porous medium,'' J. Sound Vib. **170(2)**, 161–179 (1994).

[9] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, 1986).

[10] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

Source bearing determination from a tri-axial seismometer using Rayleigh wave propagation

S. A. Stotts,[a] R. A. Gramann, and M. S. Bennett
Applied Research Laboratories, The University of Texas at Austin, P.O. Box 8029, Austin, Texas 78713-8029

(Received 17 September 2003; revised 12 January 2004; accepted 19 February 2004)

Bearing determinations for ground vehicles have been made, using a single buried three-axis seismometer. The method, based on Rayleigh wave detection, exploits the phase difference between measured wave components to obtain vehicle bearing estimates. It is referred to as a Rayleigh wave retention method and uses both prograde and retrograde motion. Determination of the propagation direction is based upon analysis of maximum correlation values obtained by cross-correlating the vertical wave components with horizontal beam components. Theoretical simulations demonstrate the technique. Methods for obtaining direction of travel and vehicle speed estimates are also discussed. Applications of the method to real data obtained from several vehicles at different test sites are presented. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1699394]

PACS numbers: 43.28.En, 43.60.Rw, 43.60.Jn [EJS] Pages: 2003–2012

I. INTRODUCTION

Time-dependent bearing estimates are essential for localization of moving sources in both water and land environments. Ground vehicles in motion generate energy in several forms, including acoustic and seismo-acoustic propagation. Buried sensors placed near a vehicle's line of travel can detect such ground motion. Three-axis seismometers can sense several propagating wave arrivals, including bulk compressional and shear waves, as well as Rayleigh waves propagating along the air–ground interface.

Three-component seismic measurements have been recorded and investigated for the past 40–50 years, mostly in application to earthquake detection. Geophysical exploration of the underlying ground structure also benefits from three-axis measurements, with recent applications involving polarization techniques.[1] Recently, Dosso[2] has used similar techniques to localize static (explosive) sources under ice using ice-mounted geophones.

The new technique discussed here was developed and demonstrated with ground vehicles, but may also be applicable to sources moving in water under ice. In each of these systems, the receiver is located in an elastic environment, whereas the sources are moving in a fluid above or below the elastic surface. The degree of coupling between the fluid and elastic layers via the source contact or motion may, however, vary for different environments and could produce different results when used to estimate bearings for under-ice sources.

Static sources are briefly mentioned to contrast the propagation effects observed with moving sources. The emphasis here is on the detection of moving sources. The results presented in this paper were obtained for ground vehicles by a detection technique that uses the vehicle-ground coupling created via tire or track motion to obtain time-dependent bearing information for vehicles that transit through the closest point of approach (CPA) to a three-axis seismometer.

The method of bearing determination presented here exploits the 90° phase difference between the horizontal and vertical components of a Rayleigh wave to obtain the bearing of the source vehicle. Other methods that exploit the 90° phase difference between the horizontal and vertical components also were investigated during this development to assess their applicability to the current problem. These techniques can generally be classified as polarization filtering. A brief discussion of these techniques is provided for parallel comparison to the Rayleigh method presented here for completeness.

A. Other methods tested for bearing determination

The first method, a motion product seismogram technique, involves multiplying a horizontal component by a 90° phase-shifted vertical component, and then integrating this product over the signal duration.[3] This method has been applied to earthquake data. The form of the derived equations, however, assumes a source originating within an infinite solid, and thus has been found to be inappropriate for the application presented here: the results were unsuccessful when applied to the real data with a moving source.

In the second method applied to the data, a three-component Rayleigh wave filter was developed for removing the Rayleigh wave portion of a seismogram. The filter is based on predictive deconvolution, in which the vertical component trace is predicted based on the horizontal component by accounting for the 90° phase lag between the components.[4] This method constructs a correlation matrix in order to build the filter coefficients to remove the surface-layer portion of the seismogram.

A third technique, recently developed at Applied Research Laboratories, The University of Texas at Austin (ARL:UT), involves the complex power spectrum. This method has been used successfully in experiments with an active source designed to detect buried objects.[5] This is in contrast to the motion product technique, the Rayleigh removal method described in the previous paragraph, and the Rayleigh wave retention method presented in this paper, which are passive detection methods. For testing this third method,

[a]Electronic mail: stotts@arlut.utexas.edu

as in the bearing determination method to be presented in Sec. II, horizontal components are formed at different angles by appropriately summing, i.e., beamforming, the two horizontal sensor components. The complex power is calculated by Fourier-transforming the signal, removing the negative frequency components, and then inverse transforming the remainder of the signal. The resulting real part is the original signal, while the imaginary part is associated with a 90° phase-shifted component. This step is performed for both the vertical and the beamformed horizontal components. Next, the product of these complex horizontal and vertical components is formed. This product's imaginary parts project the Rayleigh components and are positive for prograde or negative for retrograde. If there is no phase difference between the vertical and beamformed horizontal signals, the imaginary part is zero, corresponding to bulk wave components. Source and receiver arrays were used to test this active detection technique.

Although useful for fixed source detection and localization, the above-mentioned techniques are not readily adaptable to the moving source problem, in part because of the complicated arrival structures produced by moving sources such as those mentioned above, and none were capable of determining bearings to moving sources. The measured signals of interest contain high levels of noise due to vehicle tire to ground coupling, referred to as ground roll. Noise determination and its separation from signals by a coherent method are not straightforward for passive detection of a moving source, and time domain integration lengths are not easily determined. For the buried object determination problem, the resolution requires an active source.

Several earthquake localization techniques have been published since this initial study but were not compared to the results of this work. These methods analyze the eigenvalue structure of a data covariance matrix to determine wave polarizations.[6–8] The techniques are applied to single-station broadband seismogram data, often obtained from long-range sources, and have as a basis the cross correlation of the data either as is or after performing a Hilbert transform.[9] The arrivals analyzed usually have a high signal to noise ratio (SNR), or are preprocessed to remove noise, and require a model to obtain estimates of the parameters of interest. Advances in geoacoustic inversions may prove useful in source localization.[7,8] Further, recent inversion work applied to ocean environments may also be adaptable to such data.[10] However, for the techniques presented in this paper complicated processing is avoided in the sense that Hilbert transforming, construction of covariance matrices, or development of model-data cost functions is not required.

The Rayleigh wave retention method presented in this paper is more efficient for the sources of interest. No model of the environment or propagation is needed. Further, to perform inversions a source spectrum for each type of vehicle is needed; however, in general the type of vehicle will not be known *a priori*. Determining whether such procedures may improve the accuracy of the method is beyond this work. These methods may, however, fail for the same reasons that the above-noted methods (which were tested on several data sets) failed. Noise subtraction is not possible given that the largest contributor to the noise is the source itself via ground roll. Identifying valid integration times for processing is also not required for the method used in this paper. Averaging estimates over several segments accounts for variations in the bearing determinations, as described in Sec. II A.

B. Rayleigh wave retention method

In the bearing localization method for moving sources presented in this paper and referred to as the Rayleigh wave retention method, the three components of the data time series are segmented and then filtered in a band centered on the low frequency peak, which corresponds to the Rayleigh wave. Next, horizontal beams are formed using the x and y components. These beams are then correlated with the vertical wave components which should have a 90° phase difference with respect to the horizontal components for a Rayleigh wave. The peak correlation for each beam angle is determined. The beam angle corresponding to the maximum of these maximum correlations identifies the vehicle bearing.

Although the Rayleigh wave retention method involves use of some techniques common to the other methods discussed in Sec. I A, it was tailored to satisfy the criteria not met by those approaches.

C. Outline

Details of the Rayleigh wave retention method of bearing determination are presented in Sec. II. In Sec. III, a simulation is presented that approximates a moving source by a sum of time-shifted fields from a discrete static source to demonstrate the method for determining a ground vehicle's bearing. Section IV presents the application of the Rayleigh wave retention method to estimate direction of travel (DOT) and vehicle speed. Results from several vehicle runs are discussed in Sec. V, along with improvements to the bearing estimates based on dual mode detection. Finally, a summary is given in Sec. VI.

II. BEARING DETERMINATIONS

A. Processing method

It is known that about 67% of the energy generated from a source at the air–ground interface of a solid half-space is propagated in the form of Rayleigh waves.[11,12] This energy is generated by the coupling of the vehicle to the road via either tires or tracks. An example of an energy spectrum from a heavy vehicle near and moving toward the CPA to the sensor appears in Fig. 1. The vertical scales on this and other curves are designated as having arbitrary units (abu), since the results have been obtained using uncalibrated data. (Note that the calibration amplitude is constant over the frequencies of interest shown here; hence, only an overall magnitude correction would be necessary for actual calibration but has no relevance to the Rayleigh wave retention technique.) The vehicle was traveling on an asphalt track overlying compacted dirt at an experimental site in southern Florida. The spectrum in Fig. 1 shows typical broadband low frequency energy that has a peak near 17 Hz. This concentration of energy is caused by the vehicle-to-road coupling. The height

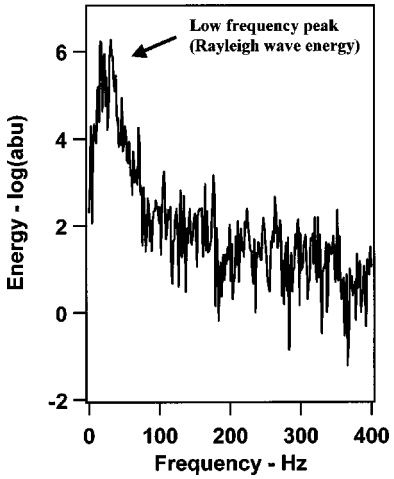


FIG. 1. Example data source energy spectrum from a moving vehicle near CPA. The data were uncalibrated.

and width of this peak are much different from those of the narrower peaks at higher frequencies. (The multiple narrow harmonic peaks, which are mostly generated by the vehicle's engine, are not discussed here.) Previous studies at ARL:UT show that the location in frequency of the broad peak is dependent on the soil environment.[13]

To begin the bearing determination procedure, the data are divided into time segments of, for example, 1024 points. The program used to determine bearing calculates the center frequency of the broad peak on a segment-by-segment basis. A relative time corresponding to each segment is recorded so that bearing angle as a function of time can be computed. Fourier transforms are then performed on each segment.

Based on this spectral analysis, the first step of the bearing determination is to filter the time series signal down to an 8 Hz band centered on the peak observed at the lower frequencies. For CPA distances greater than a few meters, most of the propagation paths are via surface waves propagating at low frequencies, usually less than approximately 25 Hz, because of attenuations in the soil; however, at this step there is still some mixture of other wave components in the data.

The second step is to form horizontal beams from the x and y components of the sensor. For each time segment the beams are steered over a set of incoming directions that might contain the signal traveling from source to receiver. The beams are given by

$$h(t)=x(t)\cos\theta_L+y(t)\sin\theta_L, \tag{1}$$

where a set of beam angles, θ_L, are constructed. The x component is assumed to be aligned parallel to the vehicle motion, i.e., parallel to the road, although this is not a requirement. Horizontal beam spacings of 10°–15° are used, since this spacing is thought to be the approximate accuracy of the bearing method; as yet, no extensive analysis has been performed to determine the optimum spacings. Each horizontal beam is then correlated with the vertical component to isolate the Rayleigh wave component of the filtered signal, based on the phase relationship between the horizontal and vertical components.

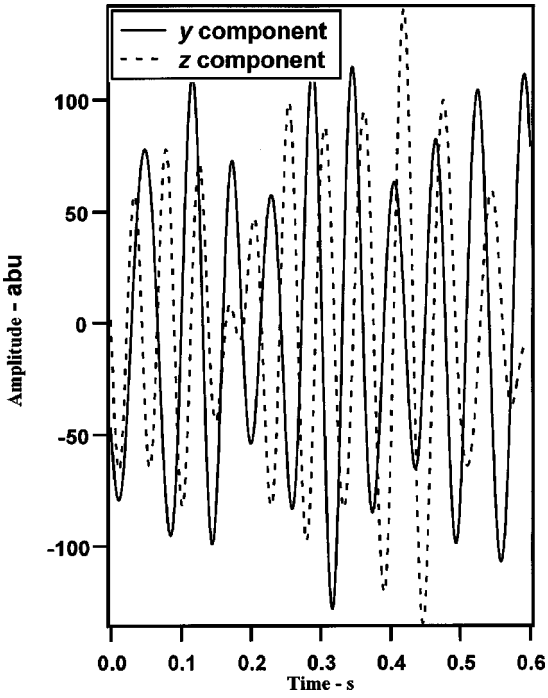


FIG. 2. Example filtered time series data. The horizontal y-component and vertical z-component phase differences are visible.

To isolate the phase relationship during the correlation, a delay time window is constructed. The window width is calculated from an assumed peak correlation given by

$$\tau_{\text{peak}}=1/(4F_c), \tag{2}$$

where F_c is the center frequency of the assumed nearly single frequency wave form components. The vertical and horizontal components are further assumed to have a 90° phase shift with respect to each other. For example, if two single frequency wave forms are constructed with these criteria and then correlated, the result as a function of delay time, is a modulated wave form with a peak correlation at the delay time.

For the data, F_c is the peak of the low frequency spectrum, as illustrated with the spectra shown in Fig. 1. The τ window for the peak search is given by

$$\tau_{\text{peak}}-\delta\tau<\tau_{\text{win}}<\tau_{\text{peak}}+\delta\tau, \tag{3}$$

to allow for the finite bandwidth of the frequency filtering and partially account for Doppler effects. Here, $\delta\tau=\tau_{\text{peak}}(F_c=17\text{ Hz})/2=0.0074\text{ s}$, for a typical Rayleigh peak frequency. The window width could be dependent on the environment; however, no adjustment of this window width was necessary for the examples given in Sec. V. The peak correlation over the values of τ_{win} is found for each horizontal beam cross-correlated with the vertical component. An example of a filtered time series segment for the vertical z component and horizontal y component is shown in Fig. 2 for the data. The correlation output for this example, shown in Fig. 3, is very similar to the result which would be obtained if two single frequency wave forms were generated with a corresponding 90° phase shift between them and then correlated. The data in the present example are taken from a heavy vehicle near CPA, whose spectrum is shown in Fig. 1.

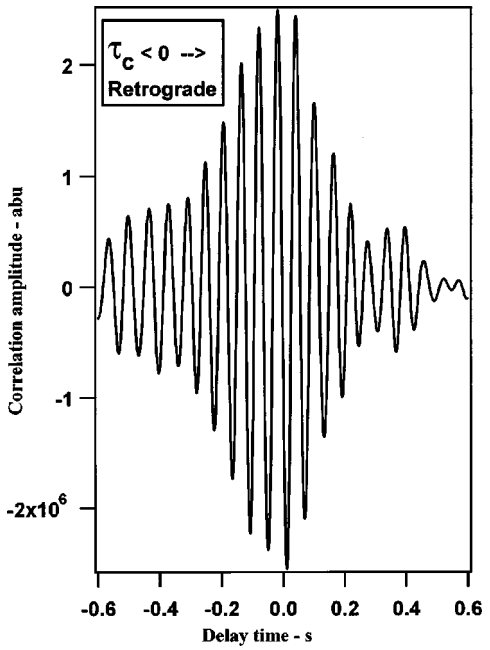


FIG. 3. Data cross-correlation for the two components plotted in Fig. 2.

The output displayed in Fig. 3 is obtained from correlating the *y* component with the *z* component. Here the maximum correlation value occurs for a negative time delay, since the *z* component lags the *y* component throughout much of the time segment. This is retrograde motion, which is expected to be the dominant mode detected for many time segments.

The next step of the procedure is to determine the beam angle that gives an overall maximum correlation within the τ window. The angle for this maximum correlation gives the bearing for the vehicle. Analysis of the hodograms for many vehicles approaching and then departing CPA showed that the amplitude of the Rayleigh wave arrivals were not constant as a function of time throughout this time period because of interference between the various wave arrivals. The hodograms displayed time periods of large Rayleigh wave motion but also showed some time segments with weak Rayleigh wave motion. This resulted in some scatter of the bearing estimates from the time segments. Averages of the maximum correlation beam angles over five time segments are performed to smooth the bearing estimates and account for these interference affects. Bearing angles outside 45° of this average were deleted from the five-segment average, and the average was recalculated.

As a final note, a simpler bearing determination was first attempted by forming beams with the *x* and *y* components, as in Rayleigh retention method, and then looking for beam with maximum output as the vehicle approached and then departed CPA. No consistent results were obtained from the vehicle at different positions along the track, i.e., this method could not track the vehicle bearing. One reason is that for these data at these sensor burial depths of a few meters, the vertical component has the most energy of all three components. The noise levels are often very high on all components except near CPA. This is mostly due to the large ground roll. This fact, combined with attenuations in the soil, and multiple arrivals from a continuously moving source, negates any ability, thus far, in extracting out the bearing by using only a horizontal beamformed $x-y$ component and simply looking for the maximum energy output. The noise shows up on most of the beams at every time step. This method was tried on several data sets and failed.

B. Retrograde and prograde detection

Often both retrograde and prograde modes are detected at the sensors. The use of two separate τ windows allows both retrograde ($\tau_c<0$) and prograde ($\tau_c>0$) motion to be filtered out separately and compared to give increased confidence in the bearing estimates. From hodograms of vertical versus horizontal components with time as an independent variable, the particle motions and directions are easily determined.

Preliminary investigations showed that good bearings may be obtained using the retrograde filter because of the dominance of retrograde motion in several of the data sets considered here. However, both types of motion are known to be strongly present for many of the time segments in these and other data sets examined. Thus, both types of motion should be considered to give a final bearing estimate in a given time segment. The final bearing values still represent only single mode estimates in environments where only a single mode is detected above the threshold.

As mentioned in Sec. II A, the initial bearing estimate was obtained by averaging every five segments, deleting bearing angles outside 45° of this average, and recalculating the average. This step was performed on both the retrograde and prograde estimates. The initial bearing estimates are then given by

$$\theta_{B,j}(\mathrm{Avg}) = (1/N_j)\Sigma\,\theta_{B,j}\,, \tag{4}$$

where the subscript *B* stands for bearing estimate, *j* is either pro (for prograde) or ret (for retrograde), and N_j is the number of valid bearing determinations that are summed out of the original five-segment averaging period. The phase difference between the two modes is taken into account before averaging by subtracting the prograde bearing angles from 180°, giving

$$\theta_{B,\mathrm{pro}} = 180° - \theta_{\mathrm{pro}}\,. \tag{5}$$

Also, confidence values were assigned for the initial bearing estimates based on the deviations from this average.

Application of this dual mode detection method to data will be presented in Sec. V B, along with more details of the processing performed to obtain accurate bearing estimates from it.

III. SIMULATIONS

A moving source was simulated to demonstrate the Rayleigh retention method. To begin, consider a static point source located in a constant sound speed, fluid half-space with a receiver in a constant sound speed, solid layer that is in contact with the fluid half-space. The solid layer is bounded at its other interface by an impenetrable vacuum half-space. This solid can be considered to be either an ice layer over water with the source in the water or, similar to

the data considered here, a ground layer below air, with the source located in the air above the solid layer.

The displacements in the solid are given by

$$u_1 = \int_0^\infty \Big[-kA_1^- e^{-\alpha_1 z} - kA_1^+ e^{\alpha_1 z} - \frac{1}{i}\beta e^{\beta z} B_1^+ + \frac{1}{i}\beta e^{-\beta z} B_1^- \Big] J_1(kr) k\,dk, \tag{6}$$

$$w_1 = \int_0^\infty \Big[-\alpha_1 A_1^- e^{-\alpha_1 z} + \alpha_1 A_1^+ e^{\alpha_1 z} + \frac{1}{i} k e^{\beta z} B_1^+ + \frac{1}{i} k e^{-\beta z} B_1^- \Big] J_0(kr) k\,dk, \tag{7}$$

where the subscript 1 refers to the solutions in the solid layer, and k is the horizontal wave number; $\alpha_1^2 = k^2 - k_{p_1}^2$ and $\beta^2 = k^2 - k_{s_1}^2$, with $k_{p_1} = \omega/c_{p_1}$ and $k_{s_1} = \omega/c_{s_1}$.[14,15] The four coefficients of the exponential terms in Eqs. (6) and (7) are obtained from the boundary conditions along with the three additional amplitudes needed to describe the up- and downgoing component fields in the fluid between the source and fluid–solid interface and the downgoing field component below the source. The construction of the analytic solution to this model is discussed in Ref. 14. Note that for a thin layer, the solutions in the solid can be written in terms of symmetric and antisymmetric modes and are useful in determining the particle motion.[16,17] The phase differences between the vertical and horizontal displacements are caused by their dependence on the Bessel functions $J_0(kr)$ and $J_1(kr)$, respectively.

The simulation assumes a vehicle traveling on an environment consisting of a single layer overlying a substrate. A full wave acousto-elastic model, OASES,[15] was used to generate vertical and horizontal component time series. The simulations were intended to be simple yet illustrate the bearing determination procedure.

The full-wave model was used to simulate a basic moving source. A comprehensive model of a moving vehicle source would require the use of many variables, some of which would pertain to the vehicle and others, specific to the environment. The environmental propagation parameters are site-specific. For the demonstrations presented here, a generic set was chosen to model a single solid layer, 20 m thick with a density of 1.5 g/cm^3, overlying a solid substrate. The compressional and shear sound speeds in the solid layer were 600 and 200 m/s, respectively. Attenuations were set to 0.5 dB/λ and 0.3 dB/λ for the compressional and shear waves, respectively, where λ is the wavelength. The substrate parameters were: sound speeds—$C_p = 900$ m/s, $C_s = 500$ m/s; density—$\rho = 2.5$ g/cm^3; and attenuations—$\kappa_p = 0.6$ dB/λ and $\kappa_s = 0.4$ dB/λ. The subscripts p and s refer to compressional and shear waves, respectively.

For the source description, a simple vertical force was used to model the features of the arrival structure. The source was located at the air–solid interface, while the buried receiver was located at a depth of 0.9 m in the solid. Using OASES, the time series were generated from this source at discrete range increments of 3 m. To simulate the moving source, a vehicle speed of 10 km/h was assumed. The initial range was at CPA = 50 m, and the maximum range was 92 m, with 14 range steps chosen for the simulation. The source time series were constructed so that most of the energy was in the band 0–50 Hz. The time series from each individual received pulse were summed. This summed time series provides a discrete approximation of the data time series obtained from the vehicle motion.

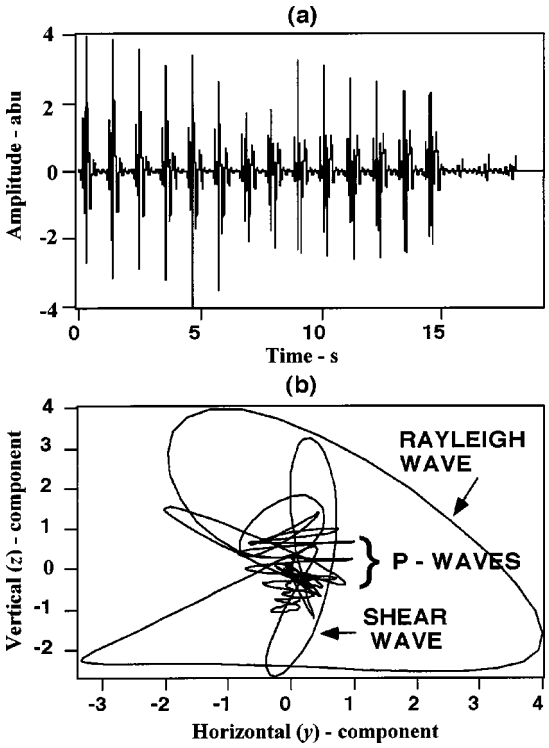


FIG. 4. (a) Simulated source vertical component. (b) Model hodogram taken near closest point of approach (CPA).

The vertical component model time series is shown in Fig. 4(a). The successive peaks are the initial arrivals of each separate point source time series. Though this replaces a continuous source with coarsely sampled discrete sources, it can still demonstrate the usefulness of the Rayleigh wave retention method at the discrete source points along the track. The modulation of the arrival amplitudes in the model time series shown in Fig. 4(a) is similar to that of the measured time series data. The data time series arrival structure will be discussed in Sec. V.

A hodogram of this model time series reveals the complicated mixture of wave types represented in the simulated data. The hodogram from the first impulse at CPA, shown in Fig. 4(b), contains compressional, shear, and Rayleigh wave arrivals. The largest arrivals are the Rayleigh waves. This structure was observed for each sampled sequence of the model time series.

After the x and y displacement components were constructed using the horizontal time series, horizontal beams were formed for each data segment using the equation given in Eq. (1) and correlated with the vertical displacement components. The maximum correlations for each horizontal beam direction were calculated for each range step (Fig. 5). The range values for the curves are designated in Fig. 5 by

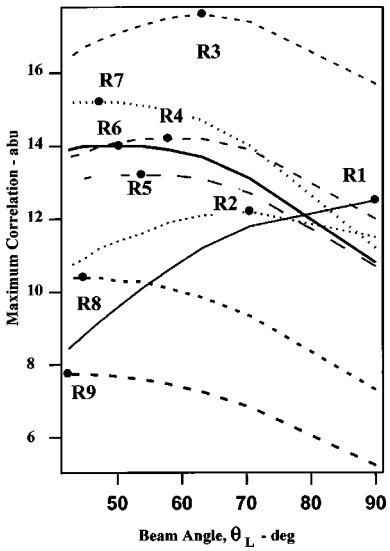


FIG. 5. Maximum correlations for simulated data for each bearing angle. The maximum of these maximum correlation values are designated by the "dots."

the labels R1, R2, etc., representing the source positions at approximately 1 s time intervals. The peaks of the maximum correlations are designated by the closed circles on the curves. As expected, the bearing angles measured for these points result in a good track of the simulated source bearing as shown in Fig. 6(a). The *z* component spectrum at R1, the 50 m range, is shown in Fig. 6(b). Similar *z* component spectra were also obtained at other ranges. The shape of the spectrum in Fig. 6(b) is broader than the data spectra in Sec. III. Again, the purpose of the simulations was to demonstrate the bearing technique, rather than exact modeling of a vehicle source.

Doppler effects of the source motion are not important in the data and thus not included in the model simulations. This is due to the low frequency of the peaks of the dominant Rayleigh wave energy. The vehicle speed, even at high speeds near 65 mph, produces a maximum Doppler shift of 2–3 Hz for the total track run, beginning from approaching

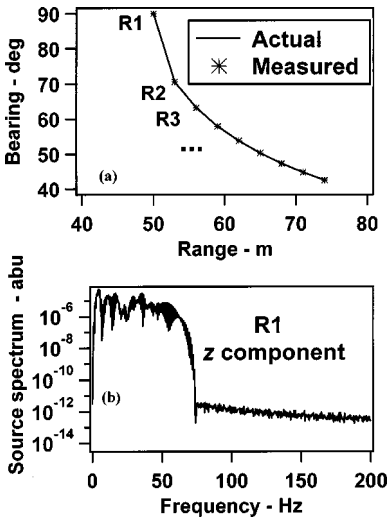


FIG. 6. (a) Bearing estimate for simulated moving source. (b) Simulated source energy spectrum.

the CPA to departing the CPA. This is on the order of one to two frequency bins for the 1488.1 Hz experimental sampling rate and 1024 point transform size. The use of a τ window and, for the data, a spectral peak that is allowed to move at each data segment accounts for any Doppler seen in the spectra.

The following points must be kept in mind when considering discrepancies in τ values between the simulated and actual data time series results presented in Sec. V. In the simulations a source spectrum slightly broader than a typical data spectrum was used [Fig. 6(b)], and the time series was not filtered at each time step. Also, a τ window was not designated *a priori* in the program for the simulated data. The Rayleigh waves' center frequency was determined by noting the τ value corresponding to the maximum correlation for each segment of simulated data. At most ranges a low Rayleigh wave frequency was found, corresponding to retrograde motion near 1 Hz. However, range values R3 and R4 were dominated by a prograde motion. The fundamental frequencies at R8 and R9 were determined to be 12.5 and 3.1 Hz, respectively. The relatively broad source spectrum in the simulations may account for variations, with respect to frequency, in the maximum correlation values at some of the ranges. An interesting point is that both prograde and retrograde motion occurs in the model and in the data.

IV. ESTIMATION OF DIRECTION AND SPEED

The estimation of the vehicle DOT and its speed are also parameters needed in localization. Further, the speed determinations are useful for vehicle classification. The estimation methods used here assume a straight-line vehicle track with the sensor placed so that the *x* axis is parallel to the source track, and the distance from the receiver to the track is assumed to be known. For actual applications the sensor orientation and its distance from the road must be determined independently. The purpose of this paper is to demonstrate how the procedure works. The methods discussed in this section can be generalized to account for other vehicle traversal geometries.

An example track is sketched in Fig. 7(a), along with an expected bearing plot, Fig. 7(b), for a vehicle approaching CPA from the left and then proceding past CPA to the right. This curve was generated by assuming constant sound speed propagation. The DOT was determined by monitoring the slope of the bearing curve as the vehicle passed CPA. A reference point was taken by averaging the first three bearing points, taken far from CPA, and then calculating the slopes of the lines from the reference to each successive bearing estimate. This line is illustrated in Fig. 7(b), where the vehicle traveling from left to right gives a negative slope. Section V contains several vehicle examples where the accumulated slope values are given.

The method for determining vehicle speed does not depend on DOT. Two estimates for the bearing are obtained, one before and one after CPA, to estimate the speed. For example, θ_1 and θ_2 shown in Fig. 7(a) for a vehicle moving from left to right. The formula to obtain speed only depends on distances with respect to CPA. Two angles are defined in

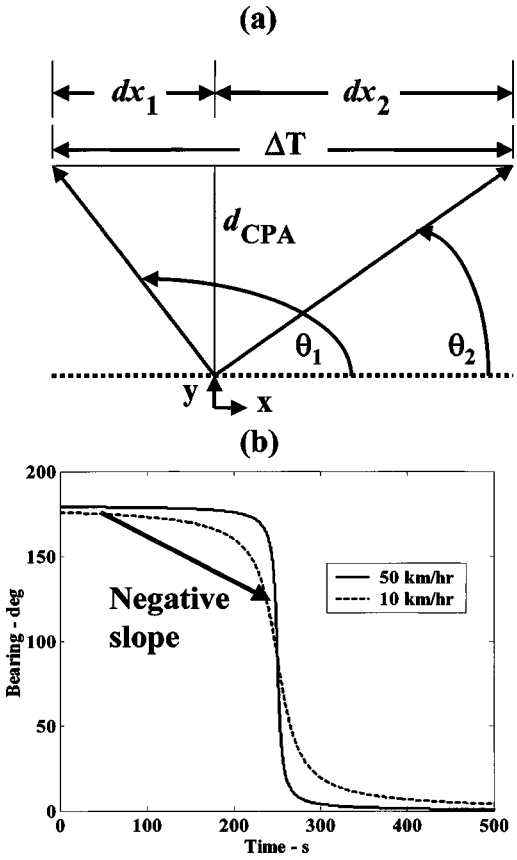


FIG. 7. (a) Experimental orientation of the sensor and quantities needed to obtain speed estimates. (b) Theoretical bearing for two speeds. Direction-of-travel (DOT) arrow is also shown to demonstrate that the slope of the arrow can determine DOT.

terms of θ_1 and θ_2, relative to an axis perpendicular to the track and passing through the sensor's y axis. These angles are defined by

$$\gamma_j = 90° - \theta_j, \quad j = 1,2. \tag{8}$$

The position along the track corresponding to θ_j is given by

$$dx_j = d_{\mathrm{CPA}} \tan(\gamma_j), \quad j = 1,2. \tag{9}$$

From the geometry, the formula for the vehicle speed is

$$v = |dx_2 - dx_1| / \Delta T, \tag{10}$$

where again the angles used to obtain dx_1 and dx_2 are the bearing estimates obtained before and after CPA, independent of the DOT.

Since the most accurate bearing estimates are obtained near the CPA, the two angles to be used in the speed estimates are those obtained closest to CPA. A more robust method would use more bearing angles in the speed determinations. In the data presented in this paper, values were obtained as proof-of-concept estimates.

For several vehicle runs presented in Sec. V, an estimated vehicle speed was noted. An expected bearing curve for these examples is included for purposes of comparison with the bearings estimated from the sensor data.

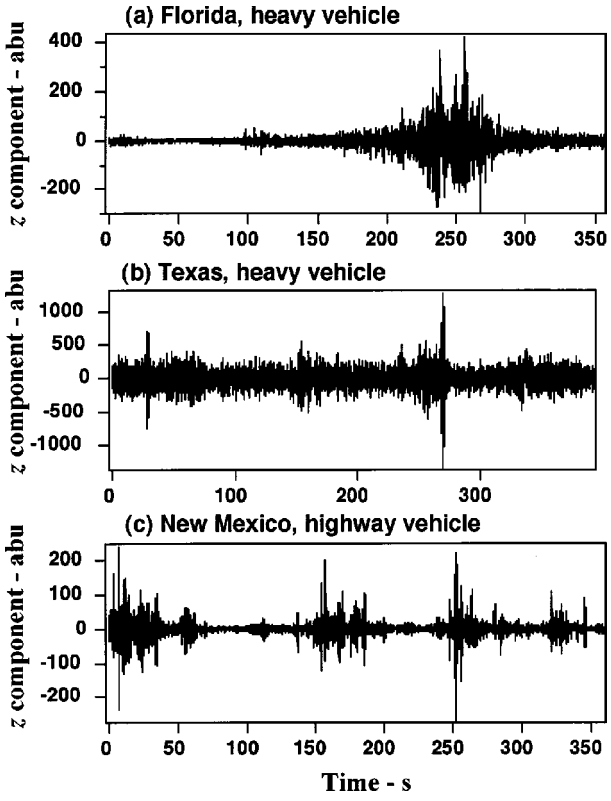


FIG. 8. Example time series from three different test sites: Florida, Texas, and New Mexico, using different vehicles.

V. BEARING EXAMPLES

In this section, several vehicle runs from different experimental sites will be presented to illustrate the Rayleigh wave retention method. In Sec. V A, only the retrograde results are presented. Vehicle bearings are estimated for all the data, and for the first site the vehicle speed is estimated and compared to the recorded vehicle speed. Section V B discusses the incorporation of both the prograde and retrograde motion into the Rayleigh wave retention method. Examples of bearing estimates are presented using both motion types. Sound speed estimates and DOT results are also presented.

A. Retrograde motion detection

Data from three different experimental sites—Florida, Texas, and New Mexico—will be presented. The Florida and New Mexico sites were in remote areas, whereas the Texas site was in an urban complex. A detailed analysis of the propagation effects at each site is given in Ref. 18, where impulsive sources were used to characterize the test fields. Some differences in the propagation for each site will be provided here as relevant to vehicle detection.

1. Florida

The first example discussed is from the experimental site in Florida. The environment consisted of an asphalt track overlying compacted dirt. Velocity dispersion curves revealed several modes of propagation, which were manifested in the bearing detection as both prograde and retrograde motion, with the dominant motion being the retrograde. An example z component time series for the heavy vehicle run at this site is shown in Fig. 8(a). This produced the spectrum in

TABLE I. Low frequency peaks and approximate widths.

Site	Peak frequency (Hz)	Peak frequency bounds (Hz)
FL	17.4	8.7–26.0
TX	4.5	1.5–7.30
NM	11.5	2.9–27.6

Fig. 1. The time series amplitude increased, then decreased as the vehicle approached CPA, and then increased again as the vehicle departed CPA. This fluctuation is caused by propagation interference effects canceling the arrivals near CPA, and by source directionality, which confines the energy more toward the line of direction of the vehicle motion. This is common behavior at other sites. A hodogram for this time series revealed regions of retrograde motion, which were identified by making a time-lapsed hodogram of the *z* vs *y* components.

The low frequency peak was at approximately 17 Hz throughout much of this run, so that the time series was filtered from 13 to 21 Hz for this data set. Table 1 summarizes the frequency bands processed from each experimental site, including the determined peak and an estimate of the bandwidth of the broad peak. For the Texas and New Mexico sites, the time series' spectra were similar to that shown in Fig. 1, except that the peaks occurred at lower frequencies, as listed in Table I.

Figure 9 shows an example of a bearing plot. The vehicle was traveling left to right along the asphalt track, with the CPA distance to the sensor being about 50 m. As shown in Fig. 9, the best estimates for bearing were obtained near CPA at a bearing of 90°. The expected bearing curve was calculated using data from the run logs. The actual vehicle speed ($V_{\log}$) is included in Fig. 9 along with the estimated vehicle speed from the data (V).

The successful bearing estimates coincided with high signal levels. To transform this observation into a threshold detector, a quantity similar to a SNR was estimated during these time segments. The absolute value squared of each data point in the *z*-component time series, such as those in Fig. 8, was taken. These values were averaged for each segment.

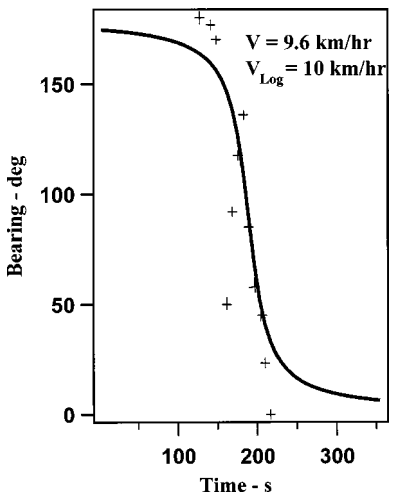


FIG. 9. Bearing for a heavy vehicle at the Florida site: estimates (+), log (−).

The first few time segments before the vehicle is detected, averaged in this way, were used as the noise floor estimate. Subsequent averages were compared to the noise estimate. A threshold limit with arbitrary units was implemented to avoid including low SNR segments. The threshold specifies the signal's averaged amplitude squared value above the noise estimate. An examination of several vehicle runs at the Florida site showed that the threshold value could be set at 10.0; this value is somewhat arbitrary and has been lowered to near 1.0 for other sites. The noise estimate was continuously updated during periods of no vehicle detection by choosing the minimum of the data averages over several time segments.

A second run of the first vehicle was also analyzed with the direction of travel being right to left. Again, good bearing estimates were obtained, which illustrates that, although vehicle detection may have some source aspect dependence, not much dependence was evident in the bearing determinations. This was probably due to the source mechanisms, either wheels or tracks, that are coupled to the ground and are symmetric with respect to travel direction.

2. Texas

Other vehicle data were obtained at a central Texas site, where several sensors were placed about 65 m from a road. The environment at the Texas site was very different from that at the Florida site. It typically consisted of approximately 1 m of packed dirt overlying an approximately 30 m layer of hard limestone. Vehicle ground excitation was dominated by Rayleigh waves, which were excited at the dirt–limestone interface. The environment was very noisy, as shown in Fig. 8(b), where an example *z*-component time series for the Texas site is depicted. SNRs obtained by calculating the ratio of the signal data (averaged amplitude squared) over the noise estimate quantify the high noise at this site as compared to the other sites. The high noise levels at the Texas site were caused by its location in an urban industrial complex with a railway and other plant facilities located nearby. This environment tests the bearing technique in a potentially high traffic urban area. Also, a hodogram produced for the time series shown in Fig. 8(b) appeared to have many wave arrivals, similar to Fig. 4(b). This was expected, given the high noise levels in the time series obtained at this site.

An example of detected bearing during a heavy vehicle traversal at the Texas site is shown in Fig. 10, along with the expected vehicle bearing. There were several runs performed in both directions past the sensors. The two bearing estimates that are higher than the bearing estimates from the logs, in the time period 100–150 s in Fig. 10, may be due to other traffic in the area. Multiple bearing tracks were evident in some of the runs at the Texas site.

3. New Mexico

Another field test in which various vehicles drove past ARL:UT's seismic sensors was conducted in northeastern New Mexico. This environment consisted of an unconsolidated dirt layer about 30 m thick overlying a more consoli-

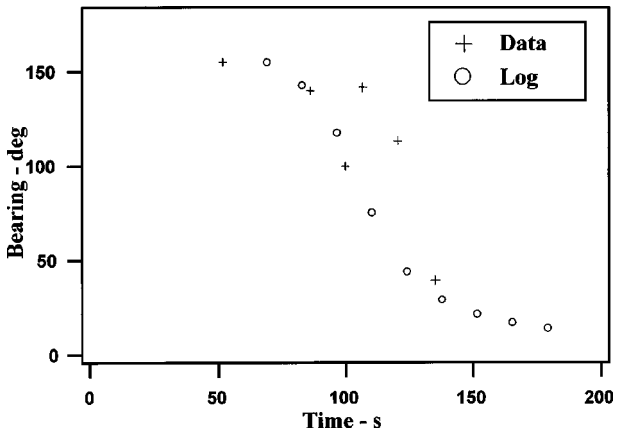


FIG. 10. Bearing for a heavy vehicle at the Texas site.

dated sediment, as determined from data analysis of thumps produced by dropping a bowling ball from a fixed height. A single dominant retrograde Rayleigh mode was excited at this site, in contrast to the Florida and Texas sites, where multiple surface modes were excited.

The experimental setup was similar to those described for the previous sites, except that several sensors were located near an intersection that contained a two-way stop. Several vehicles moving on a nearby highway were detected and these results are presented here to show that bearings can be estimated for fast moving vehicles. The *x* component of the sensor was oriented perpendicular to the highway for these vehicles.

An example time series from a vehicle traveling along a highway near this site, shown in Fig. 8(c), reveals almost discrete arrivals at certain times as the vehicle moves past the sensor. The CPA distance was estimated at several hundred yards, but was not measured. Therefore, no speed estimates were obtained for this data.

A time series segment of one of the strongest arrivals taken from Fig. 8(c) is shown in Fig. 11(a). The unfiltered hodogram of this arrival shown in Fig. 11(b) reveals the characteristic elliptical motion of Rayleigh wave propagation. Other arrivals, such as bulk compressional and shear waves, are not received because of their attenuation at these vehicle-to-sensor distances.

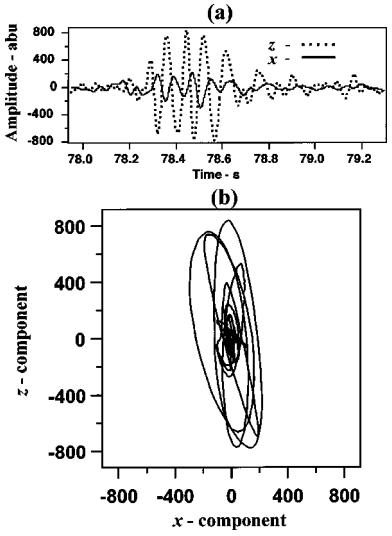


FIG. 11. New Mexico site. (a) Time series. (b) Unfiltered hodogram.

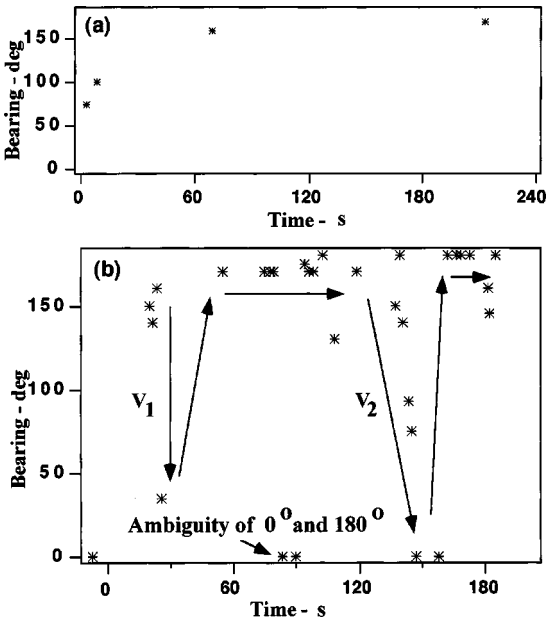


FIG. 12. New Mexico site. (a) Vehicle departing CPA. (b) Two vehicles traveling along the highway.

Since the sensor's *x* component was aligned perpendicular to the highway the bearing is expected to decrease to a minimum as the vehicle approached CPA and then increase again as the vehicle departs the CPA. The first bearing example, shown in Fig. 12(a), is for a vehicle that was first recorded as it was departing near the CPA of the sensor. Only a few bearing points were obtained because of the sparse signal arrivals for this vehicle. A better example of fast-moving vehicles is shown in Fig. 12(b), where several good points were obtained for what appear to be two separate vehicles. This example demonstrates that good bearing estimates are obtained for other orientations of the sensors horizontal components.

B. Retrograde and prograde motion detection

If both types of motion were present at a given time step, an intermediate dual-mode bearing estimate was obtained by averaging the two initial bearing estimates, provided their difference was less than 25°. This restriction gives increased confidence in the estimates when both types of modes were present but produces fewer bearing estimates.

The five-segment single-mode detections (from either type of mode) were further averaged over three consecutive single-mode initial bearing estimates. This additional averaging was done only on the initial bearing estimates that contain only one type of mode; the averaging resulted in fewer bearing estimates from single-mode detection, but increased confidence in these results. This calculation gave a single-mode intermediate bearing estimate that was compared to the dual-mode detection performed when both modes were present.

The final bearing estimate was obtained as either the dual-mode or the intermediate single-mode estimate or as an

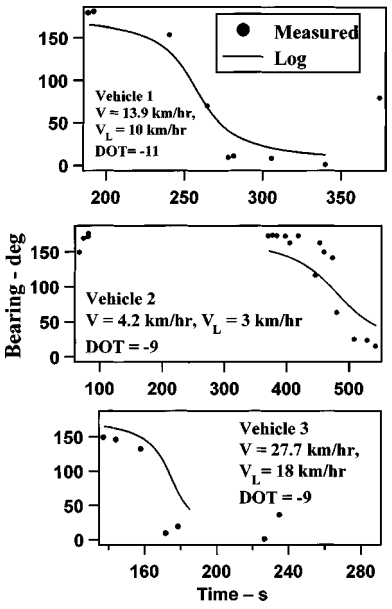


FIG. 13. Florida site bearings using dual modes, i.e., simultaneous prograde and retrograde mode detection.

average of the dual- and single-mode bearings if both intermediate values were present simultaneously over a given group of segments. The results of these final bearing estimates are shown in Fig. 13 for several vehicles at the Florida site. Vehicle speeds obtained using bearing angles near CPA are also given, along with the nominal speed noted in the run log (V_L).

Also, using the method described in Sec. IV, the DOTs were determined, and the final cumulative slope values (sum of $+1$ or -1 values representing the signs of the slopes) were calculated. These slope values were defined so that an overall minus sign indicates a 180° to 0° travel direction. These DOT values are provided in Fig. 13.

VI. SUMMARY

A method has been developed to obtain bearing estimates from moving vehicles using a single sensor to measure phase differences between horizontal and vertical components. The 90° phase shift between vertical and horizontal components for Rayleigh wave propagation was exploited to isolate the surface wave excitation created by the vehicles in order to obtain the vehicle bearing estimates. Beamforming the *x* and *y* components supplies horizontal directionality for the system. Theoretical predictions for moving sources were presented to establish the expected propagation effects from moving vehicles.

Good results were reported for three different real environments seen to have Rayleigh wave propagation in different frequency bands. Several speed estimates were made, and they compared well with speedometer readings taken during controlled test runs. Improvements to the system were demonstrated with further refinements of the Rayleigh-wave retention filtering process. Bearing estimates based on the two methods, single mode retrograde detection and simultaneous retrograde and prograde detection, were merged to provide increased confidence in the bearing estimate. The DOT values were estimated and found to correctly describe the vehicle's known path. Bearing estimates applied to other data sets have compared well to GPS tracking of vehicles during experiment runs.

ACKNOWLEDGMENTS

The authors wish to thank Ilene McCool, Dusty O'Brien, and Lewis A. Thompson of ARL:UT for many helpful discussions, and for assistance with data collection. This work was sponsored by the Defense Advanced Research Projects Agency (DARPA), with Dr. Ed Carapezza as contract monitor.

[1] E. I. Gal'perin, *The Polarization Method of Seismic Exploration* (Reidel, Dordrecht, 1984), Chap. 1.
[2] S. E. Dosso, G. J. Heard, and M. Vinnins, "Source bearing estimation in the Arctic Ocean using ice-mounted geophones," J. Acoust. Soc. Am. **112**, 2721–2734 (2002).
[3] J. E. White, "Motion product seismograms," Geophysics **29**, 288–298 (1964).
[4] J. I. Bobbitt, "Three-component Rayleigh wave filter," Oil & Gas J., 72–77 (1986).
[5] E. Smith *et al.*, "Measurement and localization of interface wave reflections from a buried target," J. Acoust. Soc. Am. **103**, 2333–2343 (1998).
[6] E. P. Chael, "An automated Rayleigh-wave detection algorithm," Bull. Seismol. Soc. Am. **87**, 157–163 (1997).
[7] C. Frohlich and J. Pulliam, "Single-station location of seismic events: A review and a plea for more research," Phys. Earth Planet. Inter. **113**, 277–291 (1999).
[8] S. Hearn and N. Hendrick, "A review of single-station time-domain polarisation analysis techniques," J. Seism. Explor. **8**, 181–202 (1999).
[9] R. G. Roberts and A. Christoffersson, "Decomposition of complex single-station three-component seismograms," Geophys. J. Int. **103**, 55–74 (1990).
[10] S. A. Stotts *et al.*, "Geoacoustic inversion in range-dependent ocean environments using a plane wave reflection coefficient approach," J. Acoust. Soc. Am. **115**, 1078–1102 (2004).
[11] L. M. Brekhovskikh, *Waves in Layered Media*, 2nd ed. (Academic, New York, 1980), Chap. 5.
[12] T. G. Gutowski and C. L. Dym, "Propagation of ground vibration: A review," J. Sound Vib. **49**, 179–193 (1976).
[13] R. A. Gramann *et al.*, "Geo-classification demonstration system description and operation guidelines," Tech. Let. TL-EV-97-24, Applied Research Laboratories, The University of Texas at Austin, 1997.
[14] S. A. Stotts, R. A. Koch, and N. R. Bedford, "Development of an arctic ray model," J. Acoust. Soc. Am. **95**, 1281–1298 (1994).
[15] H. Schmidt, "SAFARI: Seismo-Acoustic Fast Field Algorithm for Range Independent Environments: User's guide," Technical Report No. SR 113, SACLANT ASW Research Centre, La Spezia, Italy, 1987. (See also "OASES User's Guide, Version 1.7 Applications and Upgrade Notes," available from H. Schmidt, Dept. of Ocean Engineering, Massachusetts Institute of Technology.)
[16] P. Stein, "Acoustic monopole in a floating ice plate," Ph.D. dissertation, M.I.T., Woods Hole Oceanographic Institution, Cambridge, MA, 1986.
[17] P. Stein, "Interpretation of a few ice event transients," J. Acoust. Soc. Am. **83**, 617–622 (1987).
[18] S. A. Stotts, R. A. Gramann, and M. S. Bennett, "Seismic wave excitation from impulsive sources," Tech. Let. TL-EV-99-17, Applied Research Laboratories, The University of Texas at Austin, 1999.

Acoustic properties of coral sands, Waikiki, Hawaii

S. S. Fu
Hawaii Institute of Geophysics & Planetology, University of Hawaii, Honolulu, Hawaii 96822

C. Tao
Zhejiang University, Hang Zhou, 310027, China
Second Institute of Oceanography, Hang Zhou, 310012, China

M. Prasad
Department of Geophysics, Stanford University, Stanford, California 94305

R. H. Wilkens[a)]
Hawaii Institute of Geophysics & Planetology, University of Hawaii, Honolulu, Hawaii 96822

L. N. Frazer
Department of Geology & Geophysics, University of Hawaii, Honolulu, Hawaii 96822

(Received 8 May 2003; revised 18 December 2003; accepted 1 February 2004)

An *in situ* experimental study of variations of compressional wave speed and attenuation with depth in natural coral sands has been made offshore of Oahu, Hawaii. *In situ* data were collected at a center frequency of 7.5 kHz. Compressional wave speed averages around 1620 m/s and attenuation (expressed as Q_p^{-1}, the reciprocal of the quality factor) decreases from 0.04 at the seafloor to 0.01 at 2 m depth. Very little change in compressional wave speed is seen to 9 m below the seafloor. Coral sand sound speeds are lower than those reported elsewhere for quartz sand. Waveforms recorded over the upper 9 m below the seafloor exhibit virtually no peak broadening, suggesting that scattering contributes little to the *in situ* attenuation. The relationship of attenuation to frequency in the coral sands agrees with Hamilton's observations of attenuation in other sediments, although the coral sand attenuation is slightly higher than in other sediments. The coral sand relationship between attenuation and porosity also agrees with Hamilton's when the volume of intraparticle voids is deducted from the total porosity.
[DOI: 10.1121/1.1689340]

PACS numbers: 43.30.Ma [RAS] Pages: 2013–2020

I. INTRODUCTION

The geoacoustic properties of unconsolidated seafloor sediments are important factors controlling seismoacoustic propagation in the ocean. Complications arising from the variety of sediments, together with their depth and spatial variability, have presented a challenge to our understanding of acoustic behavior in marine sediments, especially in shallow water. In this study, we present the first *in situ* acoustic data from coral sand at subbottom depths greater than a few 10's of centimeters.

Coral sand deposits derived from active reefs constitute the dominant sediment cover on many modern shelves located at low latitudes in shallow, tropical to subtropical seas. They are almost ubiquitous in the many islands and atolls of the equatorial Pacific ocean. Despite the importance of coral sands in shallow water acoustics, the depth dependence of acoustic properties in natural coral sands is poorly known. Reasons for this include difficulties in coring, the usual difficulties of *in situ* measurement, and the complicated microstructure of coral sands.

We have developed an instrument, the acoustic lance, to obtain *in situ* compressional wave speed and attenuation profiles within the upper several meters of the seafloor (Fu *et al.*, 1996). This study reports lance measurements taken in a natural coral sand deposit, the Halekulani Sand Channel (HSC), off Waikiki on Oahu (Fig. 1).

Two sites were investigated in the HSC. At the first site, a 4 m long test probe deployment (Site 1 M1) was followed by a 9 m long probe deployment (Site 1 M2). At Site 2, two 2 m long probes were deployed 50 m apart (Site 2A, B). No cores were collected at Site 1, but six vibracores were recovered at Site 2 for laboratory study. We report *in situ* variations of compressional wave speed (V_p) and attenuation (Q_p^{-1}, the reciprocal of quality factor) with depth in the HSC and compare them with results of studies of quartz sands.

From *in situ* and laboratory core sample measurements, Hamilton (1980) obtained empirical relations between the acoustic and physical properties of many types of marine sediments. The parameters in those relations have been used successfully for decades to predict acoustic properties in different types of marine sediments. Lacking in Hamilton's coverage, however, were measurements from coral sand environments. Our observations of compressional wave speed between 0–10 m below the sea floor (mbsf) differ substantially from Hamilton's (1980) equation characterizing fine quartz sand.

To compare our attenuation results with the work of other investigators, Q_p^{-1} is converted to attenuation coeffi-

a)Corresponding author: Roy Wilkens;
electronic mail: rwilkens@Hawaii.edu; telephone: (808) 956-5228.

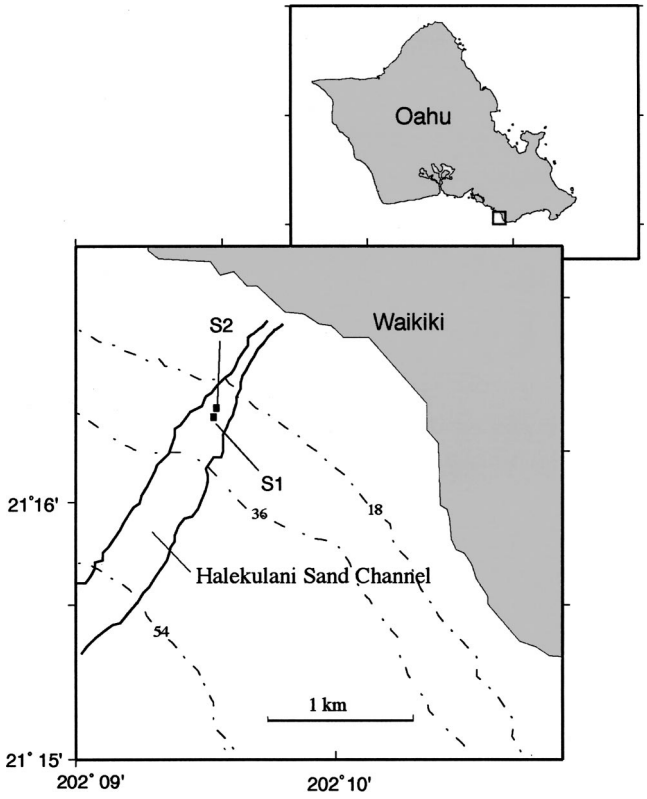


FIG. 1. Locations of acoustic lance deployment sites S1 and S2. Contour interval: 18 ft.

cients α and k. The attenuation data in coral sands agree with Hamilton's relationship between attenuation and porosity when the volume of intraparticle voids is deducted from the total porosity of the coral sands.

II. GEOLOGICAL SETTING

The Halekulani Sand Channel is located off the southern shore of the Hawaiian island of Oahu (Fig. 1). It extends from the shoreline to water depths greater than 30 m. The channel was probably part of an ancient stream drainage that was cut during a glacial low sea level stand. Sand transport in the HSC is affected by wave climate, as in the summer months when surf generated by Antarctic winter storms reaches the south shores of Oahu and reworks sands in water depths of up to 100 m. Hurricanes and tsunamis are infrequent, but their impact on sand transport is locally severe (Coulbourn *et al.*, 1988). This active environment results in a loose, poorly consolidated sediment comprising the upper 10's of meters of the sand deposits.

The HSC was chosen as the experiment site because it is a large, well-defined sand deposit. The sands show no indications of post-depositional cementation in the upper 10 m of the deposit and are poorly compacted. For this experiment, the physical parameters of the HSC sands—grain size, density, and porosity—were examined from core samples collected at the time of the experiment by Sea Engineering Inc. Samples show that the sediments within 0–8 mbsf in the channel are moderately well-sorted, greyish coral sands. A scanning electron microscope (SEM) backscattered electron image of a sample is shown in Fig. 2. The median grain size of the sand ranges from 0.19 to 0.38 mm (fine to medium sand). In general, silt and clay sizes comprise less than 10% of each sample by weight. A summary of sand characteristics in the HSC is presented in Table I. Because of wave related reworking of the sediments, there is no apparent relation between grain size and either spatial location of samples or depth of samples within the deposit.

Sediment analysis showed bulk density varying from 1.78 to 1.93 g/cm^3 and porosity varying from 0.49 to 0.57 in 39 core samples from the 6 cores recovered from the uppermost two meters of the HSC. However, the average bulk densities and porosities of each core are almost the same. Bulk densities averaged for each core are between 1.84 and 1.87 g/cm^3 and average porosities are between 0.52 and 0.54. Grain size distributions remain nearly constant in most of the samples, except for a small percentage of coral gravels present in some samples. These measurements suggest that the HSC sand deposit can be regarded as near-homogenous at meter scales (in the upper two meters), although at centimeter scales it may be inhomogeneous.

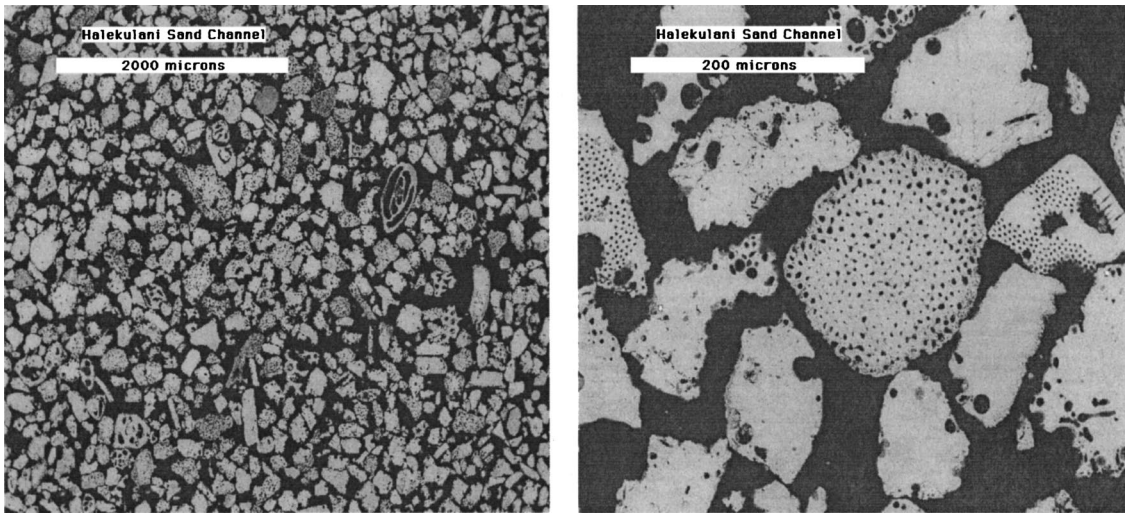


FIG. 2. Backscattered electron images of polished sections of coral sand from the Halekulani Sand Channel. Bright areas are calcium carbonate; dark areas are pore space.

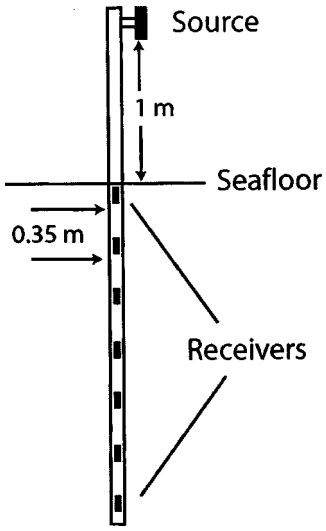


FIG. 3. Source and receiver geometry of the lance deployments at measurement Site 2. Not shown are the 2 receivers (1 m apart) above the source that were used to monitor bottom water sound speed.

III. METHODS

A. *In situ* measurement

The acoustic lance is a linear array of up to ten receivers embedded in the seafloor below an acoustic source (Fig. 3 and Fu *et al.*, 1996). It provides an *in situ* recording of full waveforms (Fig. 4) that travel vertically through the sediments. In deep-water operation a broadband acoustic source and a solid state data recording system are mounted on the weight stand of a gravity corer or an independent probe. An array of small hydrophones is mounted along the outside of the probe. In the HSC experiments, the receivers were mounted on a probe that was installed by using a water jet to penetrate the sands. After allowing a few days for sediment reconsolidation, divers brought the recording package down to the seafloor for measurements. The signal frequency band used in this analysis is 5–10 kHz, with the dominant frequency near 7.5 kHz. The spectrum of the top sediment receiver waveform in Fig. 4 is shown in Fig. 5 along with corrected spectra used for the attenuation calculation (see later).

The first deployment at the HSC (Site 1 M1) had three receivers mounted along a 4 m long probe at intervals of 2 m; all three receivers were buried in the coral sands. For the second deployment about 6 months later (Site 1 M2), 7 receivers were mounted along a 9 m subsurface probe at a spacing of 1.5 m; 2 receivers were positioned above the sediments to obtain the bottom water sound speed. The measurements were taken one week after the probe was inserted. The second deployment (Site 2A, B) was a more detailed study of acoustic structure in the upper 2 mbsf of the Halekulani sand channel. Lance measurements were made using two probes (A and B), each of which penetrated approximately two meters into the sediments; the horizontal distance between the two probes was about 50 m. Nine receivers were mounted on each probe. The top two receivers were in the water, 1 m apart, to obtain the bottom water sound speed while the remaining seven receivers penetrated the sediments at 0.35 m spacing. The first measurements were made with Probe A, two days after it was set up (Site 2A M1), and one week later a second set of measurements were made using Probe B (Site 2B M1). Two months later more data were collected from both probes (Site 2A, B M2).

Compressional wave speed (V_p) and attenuation (Q_p^{-1}) were extracted from the waveform recordings. Sound speed was estimated from the differences in signal arrival time at receivers of known separation. Errors or uncertainties in the calculations are a function of receiver spacing and the signal sampling rate (100 kHz). By interpolating the wave forms in the frequency domain, we increased the effective resolution of travel time to 2.5 μs. At a receiver separation of 35 cm (Site 2) this translates to 18 m/s in the sand—or roughly 1% of the average speed of 1620 m/s. At a separation of 1.5 m, the error is proportionally lower because the overall travel time is greater but the arrival uncertainty is constant. A complete discussion of errors and uncertainties in sound speed calculation from acoustic lance data can be found in Fu *et al.* (1996).

Effective attenuation, as it contains the effects of both intrinsic attenuation and scattering attenuation, is much more difficult to extract than sound speed, because whole wave forms must be used. Noise contamination and the correction for receiver transfer functions complicate signal processing, making a careful treatment of errors important. We corrected wave form spectra in the sediments for differences in receiver performance and for geometrical spreading by normalizing them with spectra recorded in water from the same receivers (Fig. 5). Frazer *et al.* (1999) describe in detail the two-step process used to extract attenuation (Q^{-1}) from the lance signals by an application of geophysical inverse theory to the spectral ratio method: In the first step, for every possible source–receiver interval the joint probability density function (pdf) of attenuation (effectively, spectral ratio slope)

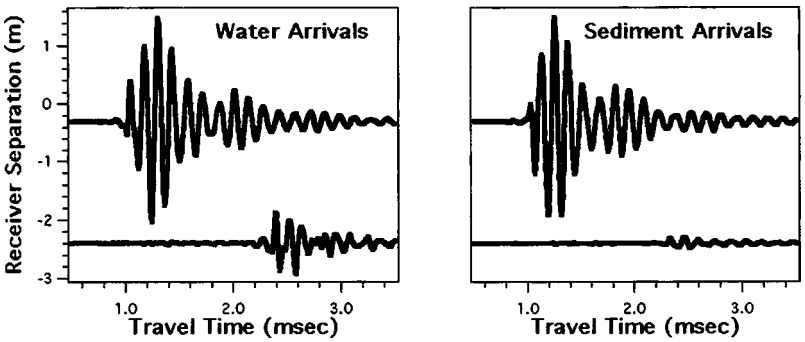


FIG. 4. Wave forms recorded by the top and bottom receivers of one of the Site 2 probes both in water (for calibration) and in the sediment. The receivers are separated by 210 cm. In each case the waveforms are normalized to the top receiver. The difference between the amplitudes of the top and bottom receivers in water (left panel) is essentially due to geometrical spreading. The greater difference in amplitudes seen in the sediment arrivals (right panel) is the result of effective attenuation.

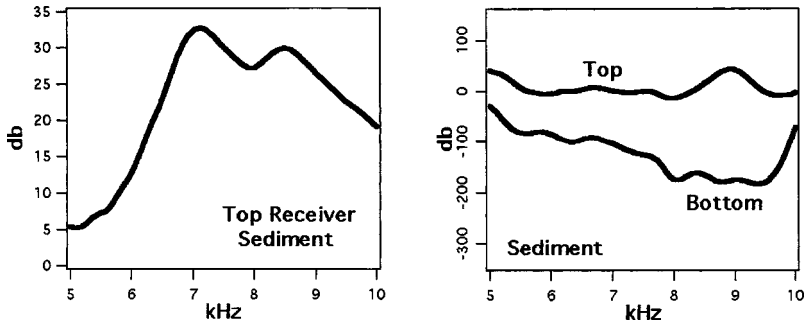


FIG. 5. Left—Spectrum of top sediment receiver waveform shown in Fig. 4. Right—Spectra for top and bottom sediment receivers shown in Fig. 4 corrected by dividing the spectra of the sediment arrivals by the spectra of the water arrivals. The difference in the slopes of these two spectra is used to calculate attenuation.

and the intercept is taken to be an L1 likelihood; a numerical integration over intercept then gives the marginal pdf of attenuation for that interval. In the second step, the peaks of these marginal pdfs are used as the data, and the half-widths of the marginal pdfs are used as the errors in an L1 likelihood taken to be the pdf of attenuation profiles. (The profile space has dimension equal to the number of receivers minus one.) Then the marginal pdf of the attenuation of each receiver–receiver (R–R) interval is obtained by numerically integrating the profile pdf over the attenuations of all other R–R intervals. A best profile is defined as the profile that agrees with the peak of the marginal pdf for each R–R interval. For brevity in this paper the marginal distributions are not shown; instead the attenuation profile uncertainty is indicated simply by plotting the envelope of the ten most probable profiles.

B. Analysis of core samples

The core samples in the HSC were collected by a vibracorer. The vibracorer uses motor-driven, counter-rotating vibrators attached to a 10 cm inner diameter, 3 m long core barrel. Samples analyzed in the laboratory were partially dry and had to be saturated with sea water under vacuum before making bulk density and porosity measurements. Vibracoring almost certainly disturbs the sediments it collects, and measurements of the porosity of core samples must be considered to be maximums. However, it must also be remembered that the HSC is located in an active wave climate and that these sands are naturally disturbed and transported repeatedly over the course of time. Thus we believe, although we cannot demonstrate, that the measured bulk properties are generally near *in situ* values.

Six samples were saturated with low viscosity epoxy, cut, polished, and examined using a scanning electron microscope (Fig. 2). Image processing of three views of each sample, at two magnifications each, was carried out to quantify the distribution of pore space between intertest and intratest pores. We used the NIH Image software package for image analysis. Pixels were first classified as either grain or pore (essentially black or white). Next, using a series of dilations followed by an equal number of erosions, the small intratest pores were removed from the image. Pixels were again classified as either pore or grain. The difference in the number of pore pixels before and after processing yielded the proportion of intratest pores—about 5% of the total, $\pm 0.5\%$. Tribble and Wilkens (1994) give a fuller description of the method, with examples.

IV. RESULTS

Waveforms recorded by the deepest probe (Site 1 M2) in the coral sands of the HSC are displayed in Fig. 6. The *in situ* compressional wave speed interval profiles from Site 1 (Fig. 7, Table II) show a speed of approximately 1595 m/s within the upper 4 m of the seafloor. The deep profile shows compressional wave speed generally increasing with depth, although that increase is not constant. Speed averages near 1600 m/s in the interval from 6.0 to 7.5 mbsf, while sound speeds above and below increase to around 1640 m/s. The long probe unfortunately could not be retrieved from the sediments for receiver transfer function calibration, which must be conducted in water (Frazer *et al.*, 1999); therefore no *in situ* attenuation data could be calculated from waveforms collected at Site 1.

It is noteworthy that little or no apparent signal broadening is seen in the self-normalized signals (Fig. 6). These wave forms are typical HSC measurements and they suggest that there is little scattering *in situ*, since scattering tends to redistribute energy from the signal onset to the coda. Thus our *in situ* attenuation calculations represent predominantly intrinsic losses.

The results of Site 2 experiments are given in Tables II and III and plotted in Figs. 8 and 9. Some of the waveforms had low signal/noise ratios due to failed receivers or to problems in recording; these waveforms were not used in subse-

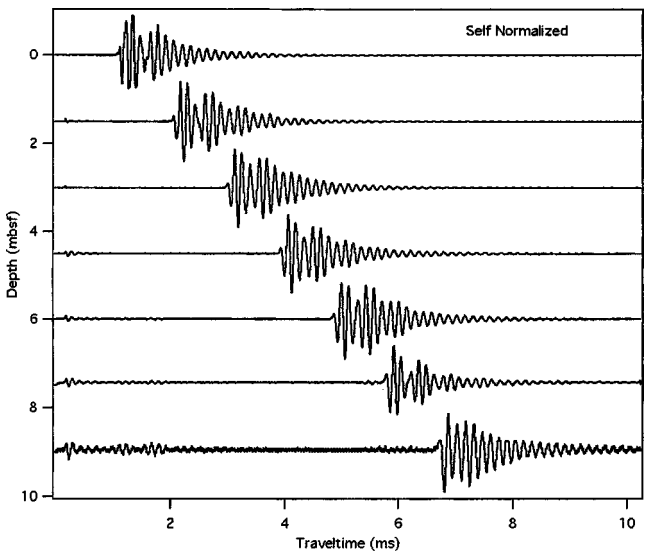


FIG. 6. Self-normalized full wave forms recorded by the acoustic lance in the Halekulani Sand Channel. Note that there is very little evidence of scattered energy arriving later in time in the more distant receivers.

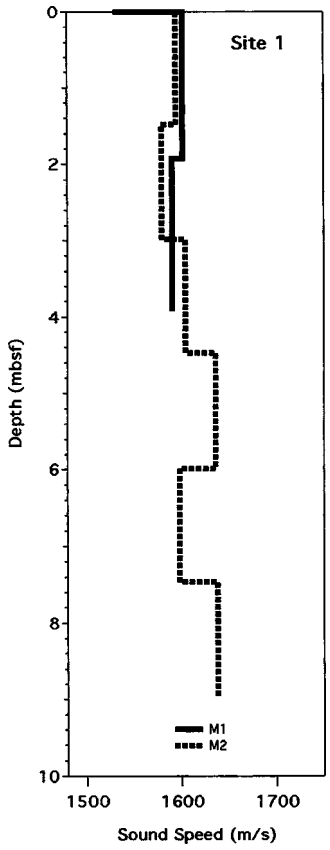


FIG. 7. *In situ* interval profiles of compressional wave speed measured at Site 1.

quent processing. At Site 2 (Fig. 8), the later measurements show some difference in compressional wave speed when compared to the measurements made two months earlier: the maximum disparity at the same depth is about 60 m/s, which is more than the expected measurement error of 1%. However, the speed inhomogeneity at the 0.35 m scale is not seen when data are averaged over 2 m. At Probes A and B the average compressional wave speeds during the second measurement over the entire 0–2 mbsf interval are virtually identical: 1631 m/s at Probe A and 1632 m/s at Probe B. The speed disparity between repeat measurements is largest at Probe A, perhaps because of sediment disturbance by the water jet. This conjecture is supported by the fact that where there are differences the later measurements are generally faster than the earlier measurements. The first measurement at Probe A was conducted only two days after it was installed, and the sediments, disturbed by the water jet, may not have completely reconsolidated in this short time. The average interval sound speeds (M1 and M2) at each of the two probes do not show an increase with depth, although interval speed varies more at Probe A than at Probe B, where it is nearly constant.

TABLE I. Summary of HSC sand characteristics.

Average median grain size	0.21 mm
Range of median grain sizes	0.19–0.25 mm
Sorting	Well sorted
Grain density	2.82 g/cm^3
Porosity average	0.53
Porosity range	0.49–0.57

TABLE II. Summary of *in situ* sound speed.

Site 1 M1		Site 1 M2	
Depth (m)	*Vp* (m/s)	Depth (m)	*Vp* (m/s)
0–1.9	1601	0–1.5	1594
1.9–3.9	1590	1.5–3.0	1579
		3.0–4.5	1604
		4.5–6.0	1636
		6.0–7.5	1598
		7.5–9.0	1638
Site 2A M1		Site 2A M2	
0–0.35	1613	0–0.35	1639
0.35–0.69	1602	0.35–0.69	1602
0.69–1.00	1672	0.69–1.00	1653
1.00–1.37	1608	1.00–1.37	1660
1.37–1.72	1613	1.37–1.72	1666
1.72–2.07	1626	1.72–2.07	1595
Site 2B M1		Site 2B M2	
0–0.68	1637	0–0.34	1602
0.68–1.37	1636	0.34–0.68	1647
1.37–1.72	1639	0.68–1.37	1636
1.72–2.07	1601	1.37–1.72	1626
		1.72–2.07	1659

Accurate laboratory measurements of attenuation at sonic frequencies (typically 1–20 kHz) are greatly limited by the sample length (the sample length must be several times larger than a wavelength). This difficulty may be addressed by performing attenuation measurements *in situ*, as with the lance. Attenuation measured from field data is termed an effective attenuation, because it includes both intrinsic and scattering attenuation. When scattering is small, as in our case, effective attenuation is approximately equal to intrinsic attenuation.

The envelopes of the ten most probable attenuation profiles for measurements 1 and 2 at Probe A and for measurement 2 at Probe B are shown in Fig. 9. For Probe B several of the wave forms recorded during the first measurement were too noisy for reliable attenuation calculations, although sound speed picks were possible. Accordingly, we show only one attenuation profile for Site 2B (Fig. 9, right panel). All three profiles show a decrease in attenuation with depth. This is most likely an effect of overburden pressure. There are three possible mechanisms for attenuation relevant to *in situ* data. One mechanism is the squirt of fluid through the solid matrix in response to deformation (Stoll, 1985; Kibblewhite, 1989). However, squirt is usually significant in high permeability solids such as sands only at low frequencies (1–1000 Hz) (Kibblewhite, 1989; Hamilton, 1980). A second possible

TABLE III. *In situ* attenuation (Q^{-1}) measurements.

Probe A	M1	M2	Probe B	Combined
Depth (m)	Q^{-1}	Q^{-1}	Depth (m)	Q^{-1}
0.00–0.69	$\geqslant$0.05	0.045	0.00–0.68	$\geqslant$0.05
0.69–1.00	$\geqslant$0.05	$\geqslant$0.05	0.68–1.37	0.025
1.00–1.37	$\leqslant$0.01	0.036	1.37–1.72	0.020
1.32–2.07	$\leqslant$0.01	0.022	1.72–2.07	$\leqslant$0.005

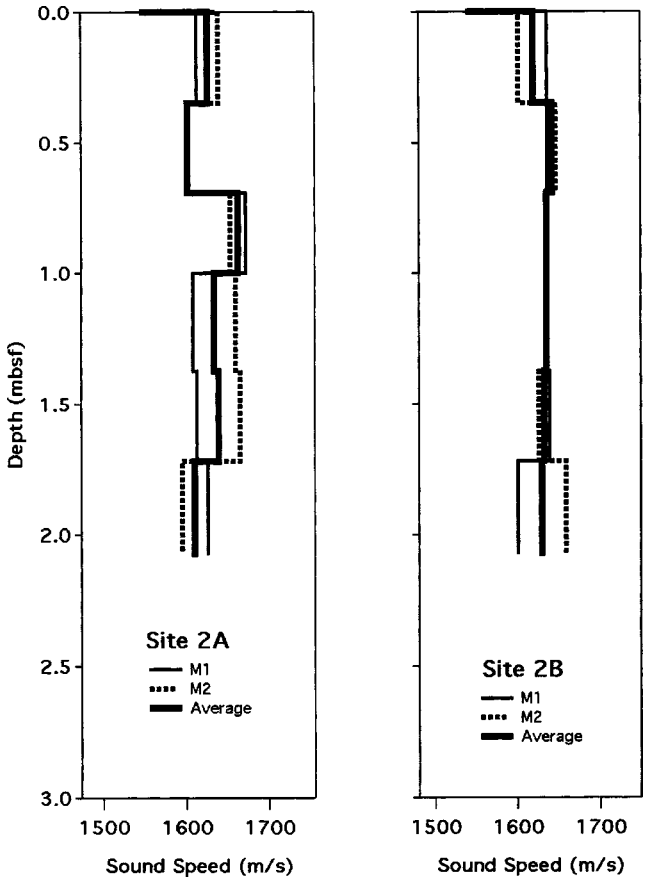


FIG. 8. *In situ* interval profiles of compressional wave speed measured at Site 2.

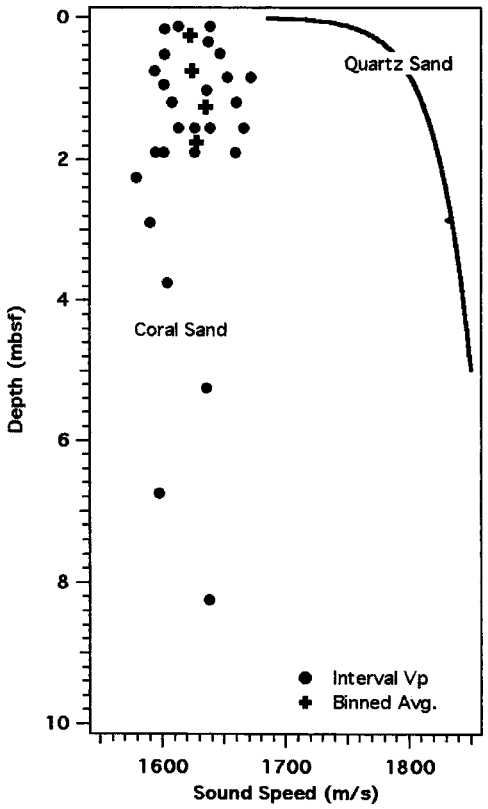


FIG. 10. *In situ* coral sand sound speed and quartz sand curve from Hamilton (1980). Points represent the middle of each measurement interval. Crosses are binned averages of all of the 2 m long probe data.

cause of attenuation is scattering due to heterogeneity such as lumps or layers. This kind of attenuation generally moves energy from the onset of the wave into the coda, a phenomenon not seen in our *in situ* data. A third possible cause of attenuation is grain boundary sliding (Stoll, 1985; Kibblewhite, 1989; Hamilton, 1972) related to low levels of compressive loading, as when contact stiffness of grains increases in deeper sediments, attenuation decreases with depth.

V. DISCUSSION

Interval sound speed data, plotted at the midpoint of each depth interval, binned averages of the shallower data, and Hamilton's (1980) empirical equation for fine quartz sand are plotted in Fig. 10. It is apparent that our *in situ* measurements do not fit speed-depth profiles for sand that are in the literature. This may well reflect the fact that much of the earlier work was based on studies of continental shelf-type quartz sands (e.g. Hamilton, 1980). Both the coral sands and the quartz sands exhibit very little increase in speed with depth over the upper 10 mbsf, but the coral sands are significantly slower than their quartz counterparts. A study of near-surface (0.05–0.30 mbsf) *in situ* compressional wave speed in carbonate sediments of the Dry Tortugas (Richardson *et al.*, 1997) yielded results generally in the same speed range as our HSC observations.

Hamilton (1976) analyzed the attenuation of compres-

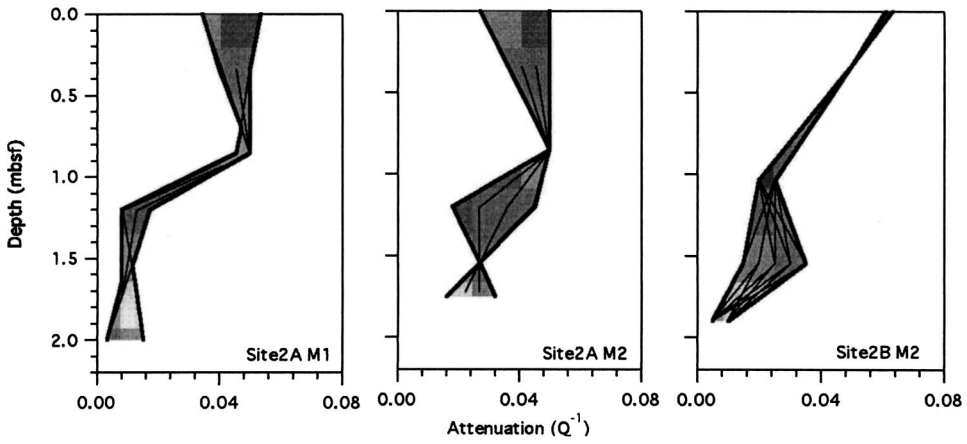


FIG. 9. *In situ* attenuation profiles from Probes A and B at Site 2. Each shaded zone is the envelope of the ten most probable profiles.

sional waves versus depth in the seafloor from various reported data at a range of frequencies. Hamilton's results were given in terms of the parameter k, which comes from the amplitude relation for an attenuating plane wave in the form

$$10\log\left|\frac{A^2(x_2)}{A^2(x_1)}\right| = -kf(x_2 - x_1), \tag{1}$$

where x_1 and x_2 are source–receiver distances, k is in dB/m kHz, and f is frequency in kHz. The relation between k and the other commonly used measures of attenuation is

$$\alpha = kf = \frac{20\pi f}{\ln(10)}\frac{1}{Qc} = \frac{8.686\pi f}{Qc}, \tag{2}$$

where α is in dB/m, c is wave speed, and the quality factor Q is dimensionless. Referring to the results of Gardner *et al.* (1964) and Hunter *et al.* (1961) on the relationship of attenuation to pressure in sands, Hamilton assumed a decrease of k with the $-1/6$ power of depth (overburden pressure) in sandy sediments. With this assumption, he computed attenuation versus depth from k values measured in surficial sediments. The results showed that, in fine sands, attenuation decreases rapidly with increasing depth to about 10 m, and then decreases less rapidly to 150 m. Our *in situ* attenuation profiles (Fig. 9) agree with Hamilton's description. Attenuation decreases rapidly in the upper 2 m interval of the HSC coral sands. Our data provide new evidence that in sandy deposits the maximum attenuation variation may take place in the uppermost several meters of the seafloor while speed itself may remain fairly constant.

Next we compare our *in situ* attenuation data for the HSC coral sand with Hamilton's (1972, 1976) regression for attenuation versus porosity, shown in Fig. 11. To minimize the experimental errors and the effects of local inhomogeneity, average attenuation values were calculated for the measurements. Our averaging scheme for *in situ* attenuation weights longer travel times—see Eqs. (6) and (24) of Frazer *et al.* (1999). The final average attenuation of the entire upper 2 m interval was taken as the mean of the three profiles ($Q^{-1} = 0.032, k = 0.53$ db/m kHz). In his regression, Hamilton used data for a variety of sediments, including clays, silts, and sands, although no pure coral sands were among these sediments. The data show that the sensitivity of attenuation to porosity varies greatly between porosities of 0.47 and 0.65 (medium sands to silt–clay). Within this porosity range, k changes from 0.52 to 0.77, then drops to 0.12. We have added an averaged *in situ* attenuation constant ($k = 0.53$) and the averaged total porosity in the HSC coral sands to Hamilton's plot as well as the sand data from Hamilton (1972) (Fig. 11). Our raw data lie on the lower boundary of Hamilton's data distribution; it seems that either k or porosity is too low for the Hamilton attenuation-porosity curves to be good predictors of attenuation in coral sand. However, total porosity includes both intraparticle and interparticle voids. Most likely, the poor agreement results from the presence of intraparticle voids in coral sands which are counted in total porosity but have little effect on the grain-to-grain contact area, a dominant factor controlling sound

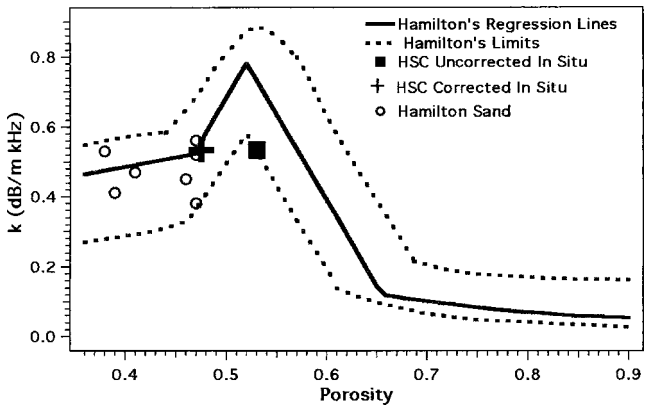


FIG. 11. Hamilton's (1976) regression lines for attenuation versus porosity. HSC *in situ* data and Hamilton's sand measurements are added to the plot. The corrected value has intragrain porosity removed.

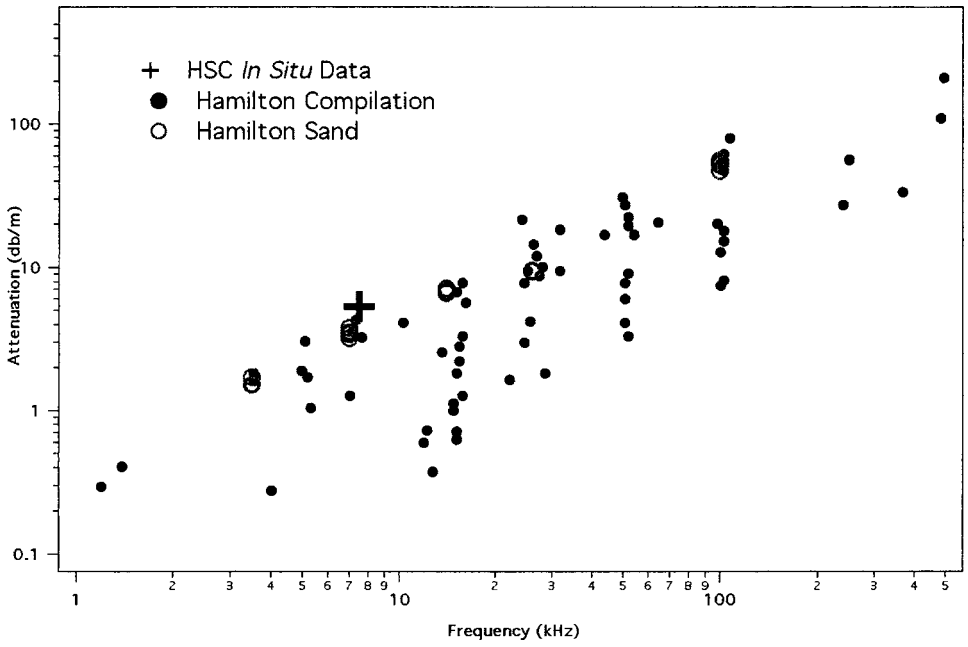


FIG. 12. Attenuation versus frequency after Hamilton (1980). Data have been added from this study and Hamilton's sand measurements have been highlighted.

wave transmission in sands (Digby, 1981; Hamilton, 1972). The results of our image processing showed that intraparticle voids occupy about 5% of the total sample volume. When the volume of intraparticle voids is deducted from total porosity, our average attenuation measurement matches Hamilton's equation well (Fig. 11). It suggests that, as Hamilton (1982) pointed out, hollow grains act as solid particles in transmitting sound waves. Because attenuation is sensitive to porosity in the sediments from medium sands to silt–clays (porosity of 0.47–0.65), our data show that caution must be exercised when using bulk porosity as an index property to predict attenuation in coral sands.

In situ attenuation is plotted in Fig. 12 as attenuation coefficient along with the data from Hamilton's (1980) compilation of all sediments. Hamilton's (1972) data are also highlighted in the plot. These coral sand average falls along the upper edge of the of the Hamilton data compilation, although it is not much greater than Hamilton's (1972) measurements made at 7 kHz.

VI. CONCLUSIONS

We have presented the first *in situ* measurements of acoustic properties in coral sand beyond a few decimeters below the seafloor. Compressional sound speeds are lower than those published in the literature for quartz sands. Attenuation is severe—some of the highest values measured in marine sediments—and decreases with depth within the upper 2 m of the seafloor. The quality of the waveforms suggests that there is very little scattering within the sediments tested, and that the attenuation measured is primarily intrinsic. Our *in situ* attenuation data agree with Hamilton's relation of attenuation to porosity if the volume of intraparticle voids is deducted from total porosity.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research Ocean Acoustics Program, by Grant No. NSFC49906004 from the National Science Foundation of China, Grant No. 2000502 from State Oceanography Administration and Grant No. 2002AA615130 from the Hi-Tech Research and Development Program of China. We also thank the staff of Sea Engineering, Inc. for allowing us to piggyback on their survey of offshore sand deposits on Oahu.

Coulbourn, W. T., Campbell, J. F., Anderson, P. N., Daugherty, P. M., Greenberg, V. A., Izuka, S. K., Lauritzen, R. A., Tsutsui, B. O., and Yan, C. (**1988**). ''Sand deposits offshore Oahu, Hawaii,'' Paci. Sci. **42**, 267–299.

Digby, P. J. (**1981**). ''The effective elastic moduli of porous granular rocks,'' J. Appl. Mech. **48**, 803–808.

Frazer, L. N., Fu, S. S., and Wilkens, R. H. (**1999**). ''Seabed sediment attenuation profiles from a movable subbottom acoustic vertical array,'' J. Acoust. Soc. Am. **106**, 120–130.

Fu, S. S., Wilkens, R. H., and Frazer, L. N. (**1996**). ''Acoustic lance: New *in situ* seafloor velocity profiles,'' J. Acoust. Soc. Am. **99**, 234–242.

Gardner, G. H. F., Wyllie, M. R. J., and Droschak, D. M. (**1964**). ''Effects of pressure and fluid saturation on the attenuation of elastic waves in sands,'' J. Pet. Technol. **16**, 189–198.

Hamilton, E. L. (**1982**). ''Sound velocity and related properties of marine sediments,'' J. Acoust. Soc. Am. **72**, 1891–1904.

Hamilton, E. L. (**1980**). ''Geoacoustic modeling of the sea floor,'' J. Acoust. Soc. Am. **68**, 1313–1340.

Hamilton, E. L. (**1976**). ''Sound attenuation as a function of depth in the sea floor,'' J. Acoust. Soc. Am. **59**, 528–535.

Hamilton, E. L. (**1972**). ''Compressional-wave attenuation in marine sediments,'' Geophysics **37**, 620–646.

Hunter, A. N., Legge, R., and Matsukawa, E. (**1961**). ''Measurements of acoustic attenuation and velocity in sand,'' Acustica **11**, 26–31.

Kibblewhite, A. C. (**1989**). ''Attenuation of sound in marine sediments: A review with emphasis on new low-frequency data,'' J. Acoust. Soc. Am. **86**, 716–738.

Richardson, M. D., Lavoie, D. L., and Briggs, K. B. (**1997**). ''Geoacoustic and physical properties of carbonate sediments of the lower Florida Keys,'' Geo-Mar. Lett. **17**, 316–324.

Stoll, R. D. (**1985**). ''Marine sediment acoustics,'' J. Acoust. Soc. Am. **77**, 1789–1799.

Tribble, J. S., and Wilkens, R. H. (**1994**). ''Microfabric of altered ash layers, ODP Leg 131, Nankai Trough,'' Clays Clay Miner. **42**, 428–436.

The calculation of radiated acoustic pressure fields from irregular multi-sided polygons

Adrian Neild[a] and David A. Hutchins
School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom

(Received 17 September 2003; revised 28 January 2004; accepted 2 February 2004)

This paper presents a method of calculating the pressure at any point emerging from a plane piston of polygonal shape in an infinite baffle. This is achieved by calculation of the impulse response as a time-stepped function, and considering the effect of each edge of the polygon in turn. The pressure is found from the convolution of this function with the time derivative of the piston velocity multiplied by the density of the propagation medium. Calculated pressure fields are shown for both a series of sources shaped as regular polygons, and for a source with an irregular shape. It is shown that the approach could be useful for the calculation of radiated fields from sources of arbitrary shape. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1690079]

PACS numbers: 43.35.Ae, 43.35.Bf, 43.20.El [JGH] Pages: 2021–2031

I. INTRODUCTION

The starting point for many calculations of transient pressure fields from transducers is the Rayleigh integral.[1] A variety of methods of producing a solution to this problem have been explored, as detailed in a comprehensive review by Harris.[2]

The calculation of pressure fields from circular pistons has been found by Stepanishen and Oberhettinger[3,4] among others, and equivalent solutions have been found for rectangular pistons.[5–7] In addition, the case of triangular pistons has been evaluated with a view to using triangles to build up more complicated shapes.[8–10] The assumptions used in these papers, and the work reported here, are that the piston is planar and fitted with an infinite rigid baffle. The propagation medium is assumed to be homogeneous, nondissipative and isotropic. Theoretical predictions have been compared in the literature to experimental measurements to show good agreement in a variety of applications. These include studies of a PVDF disk transducer operating in water at 300 kHz,[11] a 10-MHz circular hydrophone,[12] and a capacitive micromachined ultrasonic transducer (cMUT) operating in air.[13] In addition, the techniques have been further implemented to predict and compare to experimental results pressure fields from arrays, both capacitive linear arrays[14] and 2-D piezoelectric arrays,[15] and cylindrical transducers.[16,17]

In this paper, the above will be extended by evaluating the pressure arriving at any field point from an arbitrary planar polygon in a rigid baffle. This is based on the impulse response model, in which the velocity potential at a given point is found by the convolution of the source velocity with the impulse response, $h(M,t)$. The calculation of the impulse response is essentially a geometrical problem, and is found as a time-stepped function. The model that is developed here will be used to calculate the pressure fields of both a series of regular polygons of varying numbers of edges, and the field of a concave irregular piston. The technique will also be compared to existing methods. This is possible for both a square piston and an irregular concave polygon shape (which been chosen so that a solution can be obtained using four rectangles), the properties of which can be compared to the model described by San Emeterio and Ullate.[7] For the case of a 12-sided regular polygon, it will be demonstrated that the result tends towards the solution for a circle described by Stepanishen.[4] The flexibility and simplicity of the method described here is believed to be of use to workers in the field of transducer modeling and design, and in principle can be extended to transducers of any shape.

II. MODELING TECHNIQUE

A. Background and geometrical considerations

According to Huygens' principle and that of supposition a finite source of sound can be considered as an assembly of point sources. In this way, the pressure field can be modeled using a surface integral over the area of the source. There are several ways this problem can be tackled.[2] In this work, a convolution method, based on the Rayleigh surface integral, will be used, an approach sometimes called the impulse response method. Rayleigh stated that the time-dependent velocity potential, $\phi(M,t)$ is

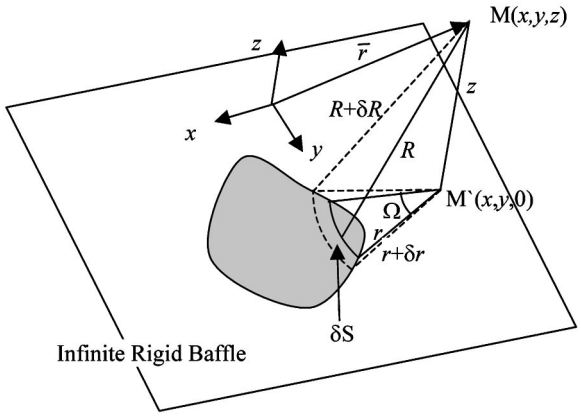


FIG. 1. Arc subtended at the projection of point P(x,y,z) onto the plane of a rectangular piston source and definitions of angles in the $z=0$ plane.

[a]Current address: Institute of Mechanical Systems, ETH Zürich, 8092 Zürich, Switzerland. Electronic mail: adrian.neild@imes.mavt.ethz.ch

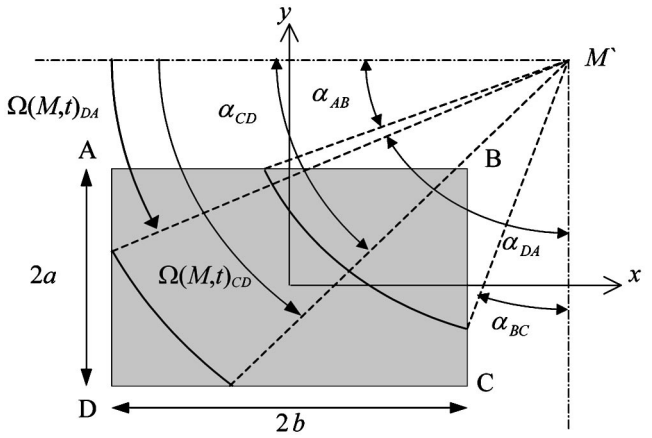


FIG. 2. Comparison of the treatment of a rectangular piston as a whole or a series of edges.

$$\phi(M,t)=\int_s \frac{\nu(t-R/c)}{2\pi R}\,dS, \tag{1}$$

where $\nu(t)$ is the normal velocity across the area of the source within the plane in which it moves, M defines the point at which ϕ occurs at time t, and R is the distance from the point M to the surface element δS (see Fig. 1).[1] The velocity potential is related to the pressure by

$$p(M,t)=\rho\frac{\partial\phi(M,t)}{\partial t}, \tag{2}$$

where ρ is the density of the acoustic transmission medium. By the use of the Dirac delta function and the convolution operator, the problem can be reexpressed as

$$p(M,t)=\rho\frac{\partial[\nu(t)*h(M,t)]}{\partial t}, \tag{3}$$

where $\nu(t)$ is the velocity of the source, and $h(M,t)$ is the scalar impulse response generated by a Dirac delta function movement of the source, which is given by

$$h(M,t)=\int_s \frac{\delta(t-R/c)}{2\pi R}\,dS. \tag{4}$$

By examining Eq. (4) and Fig. 1, it can be seen that for a given point M, the area of the source which contributes to the scalar impulse response at a given time t is the area which lies at a distance of $R=ct$ from the point. This is referred to as the sifting function of the Dirac delta pulse.[4] In effect, Eq. (4) states that the scalar impulse response is the area of the source at a distance of ct from the point M, divided by 2πR.

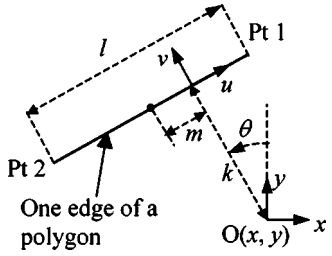


FIG. 3. Definition of terms used for each edge of a polygon.

TABLE I. Calculation of $\Omega(M,t)$ for a rectangular piston.

Time intervals	Planar time range	Resultant value of $\Omega(M,t)$
1	$t_p^B \leqslant t_p \leqslant \min(t_p^A, t_p^C)$	$\pi/2-\alpha_{AB}-\alpha_{BC}$
2	$t_p^A \leqslant t_p \leqslant t_p^C$	$\alpha_{CD}-\alpha_{AB}$
3	$t_p^C \leqslant t_p \leqslant t_p^A$	$\alpha_{DA}-\alpha_{BC}$
4	$\max(t_p^A, t_p^C) \leqslant t_p \leqslant t_p^D$	$\alpha_{CD}+\alpha_{DA}-\pi/2$

To simplify Eq. (4), consider the small area in the plane of the piston δS, as R is increased to $R+\delta R$, as shown in Fig. 1. This area can be found by considering the sphere centered at the point M. The plane of the piston will intersect this sphere, and the line of intersection will be a circle centered at $M`$. Using the coordinate shown in Fig. 1, the distance to point M is defined as R, and hence the radius of this circle is $r=(R^2-z^2)^{1/2}$. Similarly, if the distance to point M is $R+\delta R$, then $r+\delta r=[(R+\delta R)^2-z^2]^{1/2}$. This in turn means that $r\,\delta r=R\,\delta R$. The small area (δS) in Fig. 1 can be expressed as $\Omega r\,\delta r$, which equates to $\Omega R\,\delta R$. Hence, the surface integral expressed in Eq. (4) can be expressed as an integral with respect to δR over the infinite range of R to give

$$h(M,t)=\int_0^\infty \frac{\delta(t-R/c)\Omega(M,R)R}{2\pi R}\,dR, \tag{5}$$

where the angle Ω has been expressed as a function of the field point M, and the distance from this point to the surface element R. If $\tau=R/c$, then

$$h(M,t)=\int_0^\infty \frac{\delta(t-\tau)\Omega(M,\tau)c}{2\pi}\,d\tau. \tag{6}$$

As discussed earlier, the sifting function of the Dirac delta is such that this integral can be simplified to

$$h(M,t)=c\frac{\Omega(M,t)}{2\pi}, \tag{7}$$

where the angle $\Omega(M,t)$ refers to the angle subtended by the arc which lies within the piston, at point M, for a given time point t. This is a result which lends itself well to numerical modeling in which subsequent time steps are considered. In this way, the problem is reduced to one of computational geometry. The result is a time-limited function, where the limits are the shortest and longest path lengths from the point to the source.

Rather than consider the source as a whole entity in order to calculate impulse response, it is possible to consider each edge in turn; this has been described by Skudrzyk[18]

TABLE II. Calculation of $\Omega(M,t)$ for a rectangular piston considered as series of edges.

Time intervals	$\Omega(M,t)_{\text{edge}}$
$t_p^B \leqslant t_p \leqslant t_p^A$	$\Omega(M,t)_{AB}=\alpha_{AB}$
$t_p^B \leqslant t_p \leqslant t_p^C$	$\Omega(M,t)_{BC}=\pi/2-\alpha_{BC}$
$t_p^C \leqslant t_p \leqslant t_p^D$	$\Omega(M,t)_{CD}=\alpha_{CD}$
$t_p^A \leqslant t_p \leqslant t_p^D$	$\Omega(M,t)_{DA}=\pi/2-\alpha_{DA}$

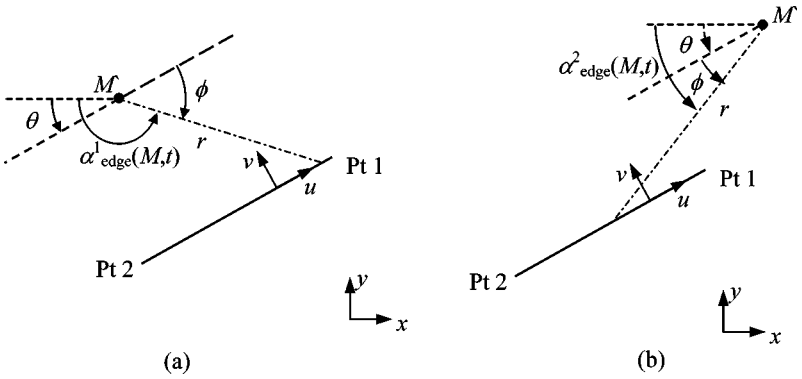


FIG. 4. Illustration of the calculation of $\alpha_{edge}(M,t)$ for (a) case 1 and (b) case 2.

among others. This has been utilized by the authors to produce a model for a rectangular source and a finite rectangular receiver in the time domain;[19] here a similar technique will be used to produce field predictions for multiple straight-sided shapes.

B. A new geometric technique

The new technique described in this section treats a source as a series of edges, and uses the geometry of these edges to calculate the angle subtended at point M' by the source at each given time point, $\Omega(M,t)$. An example will be given to show how this method compares to the solution provided by San Emeterio and Ullate for a rectangular source.[7]

First, a simplification can be introduced allowing the problem to be considered in just two dimensions. This can be achieved if the time (t) is redefined as planar time (t_p) within the x-y plane according to

$$t_p = \frac{\sqrt{c^2t^2 - z^2}}{c}. \tag{8}$$

The origin is defined as being at the center of the rectangular piston source, with four corners labeled A–D as shown, of size $2a \times 2b$. An example that illustrates the model developed by San Emeterio and Ullate[7] for the case where the point M' lies in the region outside the projection of edges AB and BC (that is for $x \geq b$ and $y \geq a$) is shown in Fig. 2. Arcs centered at $M'(x,y,0)$ for two different points in time are shown, and the radius of each arc is given by ct_p. The intersection of these arcs with the source edges can be defined using the four angles α_{AB} to α_{DA} shown. The resultant geometry can be considered more easily using four new variables, which can be defined as

$$\begin{aligned} d_1 &= y - a, \\ d_2 &= x - b, \\ d_3 &= y + a, \\ d_4 &= x + b. \end{aligned} \tag{9}$$

These variables allow the planar times at which the arc crosses each piston corner to be defined as

$$\begin{aligned} t_p^A &= \sqrt{d_4^2 + d_1^2}/c, \\ t_p^B &= \sqrt{d_1^2 + d_2^2}/c, \\ t_p^C &= \sqrt{d_2^2 + d_3^2}/c, \\ t_p^D &= \sqrt{d_3^2 + d_4^2}/c. \end{aligned} \tag{10}$$

The angles α_{AB} to α_{DA} shown in Fig. 2 can then be found as

$$\begin{aligned} \alpha_{AB} &= \sin^{-1}(d_1/ct_p), \\ \alpha_{BC} &= \sin^{-1}(d_2/ct_p), \\ \alpha_{CD} &= \sin^{-1}(d_3/ct_p), \\ \alpha_{DA} &= \sin^{-1}(d_4/ct_p). \end{aligned} \tag{11}$$

By using these angles, $\Omega(M,t)$ can be found for any point M, which lies within the region defined by $x \geq b$ and $y \geq a$, using the equations presented in Table I. The result above is the same as that of San Emeterio and Ullate,[7] except for the

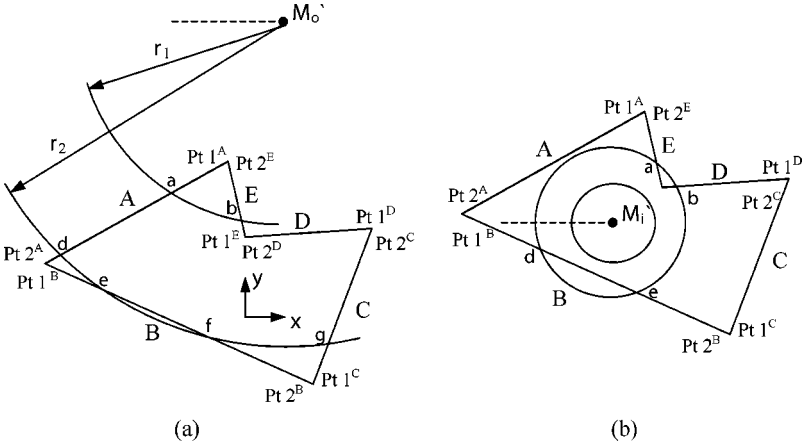


FIG. 5. Arcs crossing an irregular polygon emanating from a field point M', located either outside (M_o') or inside (M_i') the piston.

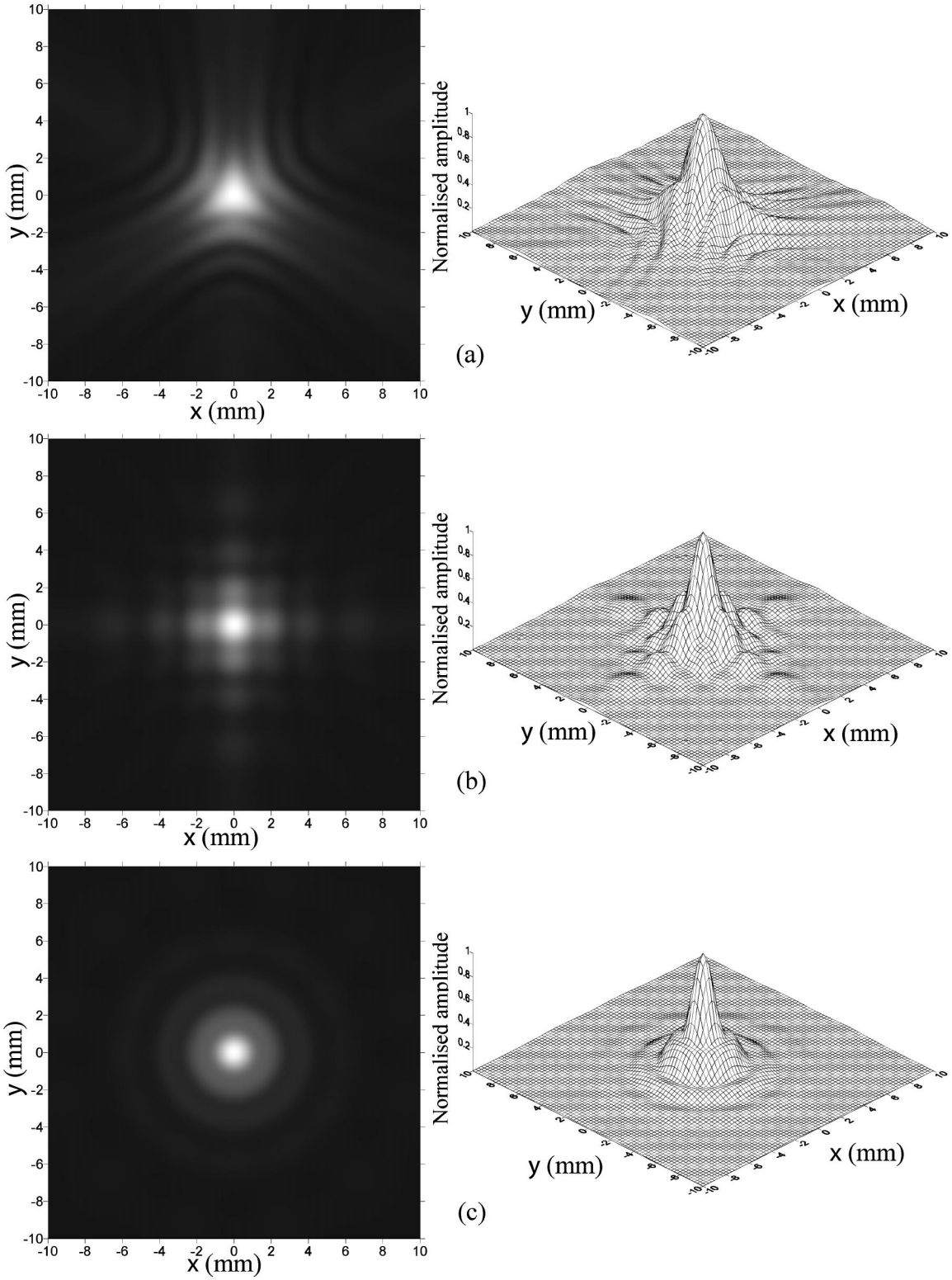


FIG. 6. The pressure field in the *xy* plane at a distance of $z=7$ mm from the source, for (a) a triangle, (b) a square and (c) a dodecagon of constant area.

redefinition of some terms and the use of planar time.

The approach described above treats the source as a whole. However, it is also possible to treat each edge separately. When such a method is used, the relationship between planar time and $\Omega(M,t)$ for any particular edge $[\Omega(M,t)_{\text{edge}}]$ can found using Table II. As an example, the angles $\Omega(M,t)_{\text{DA}}$ and $\Omega(M,t)_{\text{CD}}$ are labeled in Fig. 2. Outside the time interval defined for each edge in Table II, $\Omega(M,t)_{\text{edge}}=0$. The subtended angle at a point in time, $\Omega(M,t)$, can then be found using Eq. (12):

$$\Omega(M,t)=-\Omega(M,t)_{AB}+\Omega(M,t)_{BC}+\Omega(M,t)_{CD}-\Omega(M,t)_{DA}. \tag{12}$$

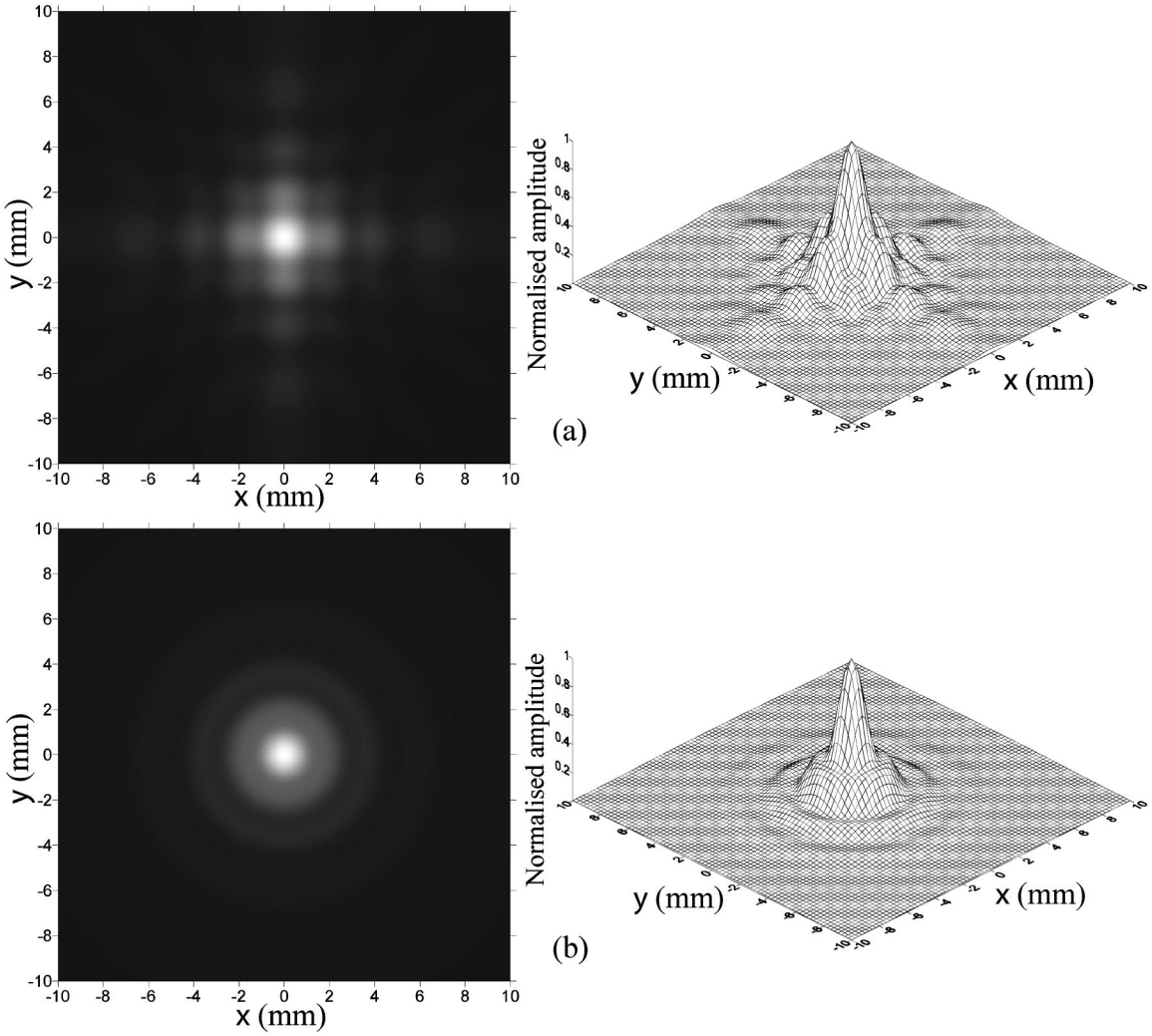


FIG. 7. The pressure field in the *xy* plane at a distance of $z=7$ mm from the source, for (a) a square and (b) a circle of constant area, using previously published techniques[4,7] for comparison to Figs. 6(b) and (c), respectively.

Although both these methods can be seen to yield the same answer, the second method can also be used to calculate the fields of more complicated piston shapes. Provided all the edges are considered, and the source is pieced together correctly, it would be possible to apply this method to any polygon. Having established the equivalence of the two approaches for the modeling of a rectangle, a more general solution for a single edge will be derived in the following section.

C. A general solution

In order to find a general solution it is necessary to define some new variables, as shown in Fig. 3 for an arbitrary edge (shown as a solid line). The two ends of the edge have been labeled Pt 1 and Pt 2. An (x,y) coordinate system is shown with an arbitrary origin; this remains fixed for all edges completing the polygon. However, a second coordinate system (u,v) is defined in which the origin shifts for each edge. If a line is drawn from the *xy* origin such that it intersects the edge (or the projection of the edge) at 90°, then this point is the origin of u, v system. The u axis lies along the line of the edge such that if Pt 1 is located at u_1 and Pt 2 at u_2, then $u_2<u_1$. The v axis is defined as being at 90° from the u axis measured anticlockwise. The angle θ is the angle of the v axis with respect to the y axis measured anticlockwise. The variables k, l, and m define the length and location of each edge and are the inputs, along with θ, to the model, which describes the shape of the piston. The distance k is the separation between the origins of the two coordinate systems measured in the direction of v, the length of the edge is defined as l, and m is the distance of the origin of the (u,v) coordinate system from the center of the edge measured in the u direction.

As indicated in the previous section, it is necessary to calculate the angle at a point M', of the arc of radius r (where $r=ct_p$), which is subtended by the acoustic source. The method shown here treats the polygon as a series of edges. For an arc of radius r the angle of this arc subtended by the source can be found from the angles at which it intersects each edge, measured from a fixed reference line. This reference line is defined as being in the negative x direction from M' with the angles measured counter-clockwise and termed $\alpha_{\text{edge}}(M,t)$.

If the point of interest, M, is located at (x_M,y_M,z_M), then u_M and v_M, are given by

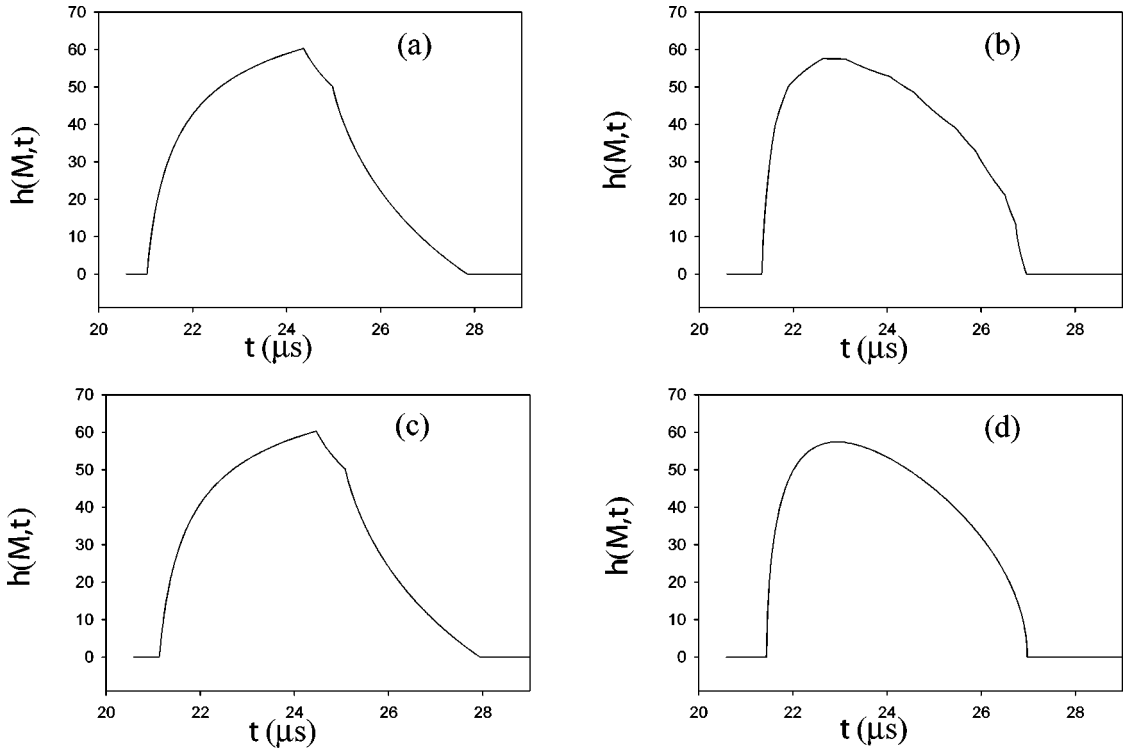


FIG. 8. The impulse response $[h(M,t)]$ calculated by the present method at a field point $x=2.5$ mm, $y=3$ mm and $z=7$ mm for (a) a square source and (b) a dodecagon. This can be compared to existing solutions for (c) a square source[7] and (d) a circular source.[4]

$$u_M = y_M \sin\theta + x_M \cos\theta,$$
$$v_M = y_M \cos\theta - x_M \sin\theta - k. \tag{13}$$

Figure 4 shows two scenarios of a line of length r meeting the edge under consideration. In Fig. 4(a), M' is located at a lower u than the point of intersection; this is termed case 1. This occurs if $u_M < m + l/2$. As the point in time varies, the length of the line r will also change. Consequently, at some time values an arc of radius r located at M' will intersect the edge, whereas at other time values no intersection will occur and hence $\alpha_{\text{edge}}(M,t)$ is set to 0. The time point at which the arc first intersects the line is termed t_p^{start}, and the time at which the last intersection occurs is termed t_p^{end}. The relevant times are given by

$$t_p^{\text{start}} = |v_M|/c \quad \text{if } u_M > m - \frac{l}{2},$$
$$t_p^{\text{start}} = \frac{\sqrt{v_M^2 + (u_M - m + l/2)^2}}{c} \quad \text{if } u_M \leqslant m - \frac{l}{2}, \tag{14}$$
$$t_p^{end} = \frac{\sqrt{v_M^2 + (u_M - m - l/2)^2}}{c}.$$

For case 1, $\alpha^1_{\text{edge}}(M,t)$ can be found using

$$\alpha^1_{\text{edge}}(M,t) = 0 \quad \text{if } t_p \leqslant t_p^{\text{start}} \quad \text{or } t_p \geqslant t_p^{end},$$
$$\alpha^1_{\text{edge}}(M,t) = \theta + \pi - \phi \quad \text{if } t_p^{\text{start}} < t_p < t_p^{end}, \tag{15}$$

where ϕ, as shown in Fig. 4(a), is given by

$$\phi = \sin^{-1}\left(\frac{v_M}{r}\right). \tag{16}$$

Case 2 occurs when M' is located at a higher u than the point of intersection. This occurs if M' is located at $u_M > m - l/2$. For case 2, t_p^{start} and t_p^{end} are given by

$$t_p^{\text{start}} = |v_M|/c \quad \text{if } u_M < m + \frac{l}{2},$$
$$t_p^{\text{start}} = \frac{\sqrt{v_M^2 + (u_M - m - l/2)^2}}{c} \quad \text{if } u_M \geqslant m + \frac{l}{2}, \tag{17}$$
$$t_p^{end} = \frac{\sqrt{v_M^2 + (u_M - m + l/2)^2}}{c}.$$

For case 2, $\alpha^2_{\text{edge}}(M,t)$ can be found using

$$\alpha^2_{\text{edge}}(M,t) = 0 \quad \text{if } t_p \leqslant t_p^{\text{start}} \quad \text{or } t_p \geqslant t_p^{end}$$
$$\alpha^2_{\text{edge}}(M,t) = \theta + \phi \quad \text{if } t_p^{\text{start}} < t_p < t_p^{end}, \tag{18}$$

as shown in Fig. 4(b), where ϕ, is given by Eq. (16).

It should be noted that for particular locations of M', and particular moments in time, both cases may occur simultaneously for the same edge.

Figure 5(a) shows an example of an arbitrary five-sided polygon. The edges of the polygon have been labeled using capital letters. For each edge the two ends are labeled Pt 1 and Pt 2, with the edge as a superscript for clarity. The labeling of the points is important, with Pt 2 of one edge being Pt 1 of the next, and the convex parts of the piston are labeled counter-clockwise from Pt 1 to Pt 2. In Fig. 5(a) a point has been labeled M_o', and from this point two arcs are drawn crossing the piston, representing two moments in time. It can be seen that for the arc of smaller radius, the angle which needs to be calculated is that from the intersect

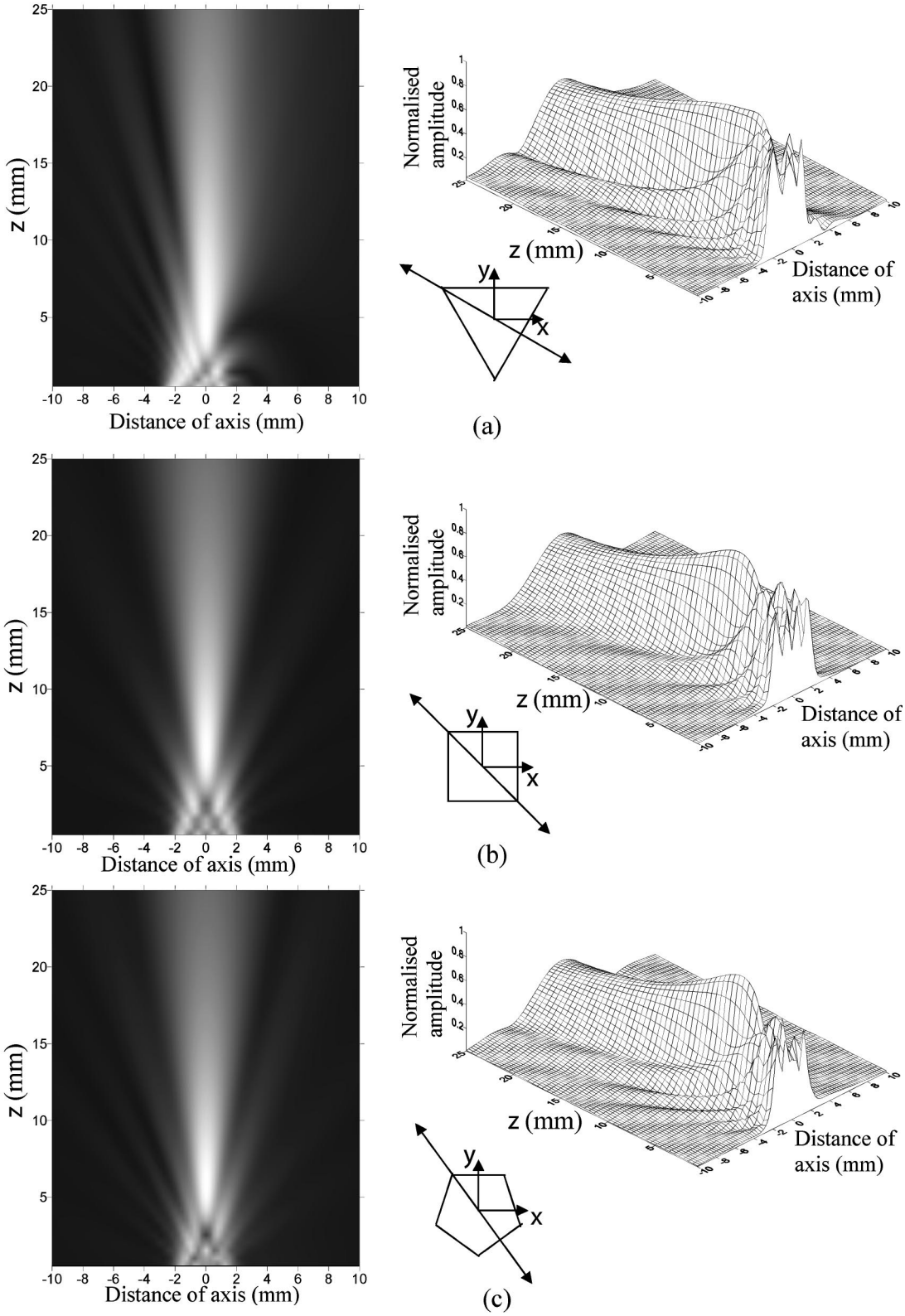


FIG. 9. The radiated pressure field in the planes shown for (a) a regular triangle, (b) a square, and (c) a pentagon of constant area.

labeled *a* to that labeled *b*. Thus, the angle subtended at $M_o^{`}$ is given by the angle to *b* minus that to *a*. These two angles are found using edge E case 1 and edge A case 2, respectively. For the larger radius arc, the required subtended angle is the angle to *g* minus that to *f*, added to the angle to *e* minus that to *d*; these would be found from edge C case 1, minus edge B case 2, plus edge B case 1, minus edge A case 2. It can be seen that in both these examples the angle at the instant moments in time depicted are found by the general rule of case 1 minus case 2. Figure 5(b) shows the same polygon, with point $M_i^{`}$ located inside the piston area; two circles have been drawn representing two time points. For

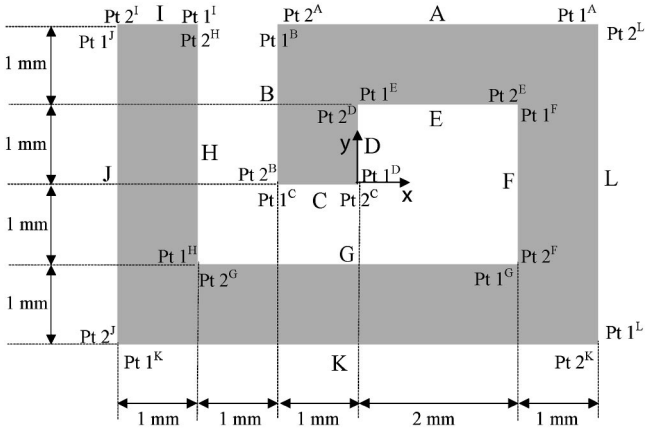


FIG. 10. The labeling and shape of an irregular polygonal piston radiator.

the larger circle, the angle subtended at $M_i^{`}$, measured counter-clockwise, can be found from the intersect *d* minus *a*, plus *b* minus *e*, which relates to edge B case 1 minus edge E case 2, plus edge D case 1 minus edge B case 2. Again the rule that the angle is found by summing the angles for each edge such that case 1 is positive and case 2 is negative holds true. For the smaller circle shown in Fig. 5(b) the angle subtended at $M^{`}$ is equal to $2\pi$; however, as at this point in time no edge is crossed, all the values for each $\alpha_{\text{edge}}$ are zero, consequently it is necessary to include a new term for when $M^{`}$ lies within the piston area, called case 3, such that

$$\alpha^3 = 0 \quad \text{if} \quad t_p \geqslant \min(t_p^{\text{start}}),$$
$$\alpha^3 = 2\pi \quad \text{if} \quad 0 < t_p < \min(t_p^{\text{start}}). \tag{19}$$

It is important to note that this case only arises if $M^{`}$ is located within the area of the piston, and is 2π up until the time when the first edge is crossed, so until the minimum value of t_p^{start} that was calculated for either case for any of the edges.

It is now possible to calculate $\Omega(M,t)$ using

$$\Omega(M,t) = \sum_{n=1:N} \alpha_n^1(M,t) - \sum_{n=1:N} \alpha_n^2(M,t) + \alpha^3(M,t) + 2C\pi, \tag{20}$$

where the edge number is represented by *n* which is from $1{:}N$, where *N* is the total number of edges, and where *C* is an integer such that $\Omega(M,t)$ lies between 0 and 2π.

III. APPLICATION TO REGULAR POLYGONS

In order to demonstrate this modeling technique, the pressure fields have been calculated from a series of regular polygons. Each of the polygons had a surface area of 12.25 mm^2, which equates to a 3.5-mm sided square. The polygons investigated have 3, 4, 5, 6, 8 and 12 sides. The simulations are performed in air, with $c = 343\ \text{ms}^{-1}$.

The input variables of the model, which define the shape of the polygon, are *N* the number of sides, and for each side *k*, *l*, *m*, and θ. The origin of the *xy* coordinate system was defined as being located at the center of each polygon, and the polygon was oriented such that θ for edge 1 was 0°, and each further edge was then labeled counter-clockwise. This allowed these variables for edge *n* (where *n* is between 1 and *N*, the total number of sides) to be calculated as

$$\theta_n = \frac{2\pi(n-1)}{N},$$
$$k_n = \sqrt{\frac{\text{Area}}{N\tan(\pi/N)}},$$
$$l_n = 2k_n\tan(\pi/N), \tag{21}$$
$$m_n = 0.$$

A criteria that can be applied to establish if $M^{`}$ lies within the piston is that this is the case if v_M is negative for all edges.

Pressure fields have been calculated in both *xy* planes and planes of varying *z*. The waveform $[v(t)]$ representing the movement of the piston used in the simulations was a toneburst of 20 cycles, with a center frequency of 500 kHz. Figure 6 shows the peak-to-peak pressure field in the *xy* plane (i.e., at a constant distance away from the piston), at a value of $z = 7$ mm, for (a) a regular triangle, (b) a square and (c) a dodecahedron. In each case, the lines of symmetry of the piston correspond to the symmetry seen in the field, as would be expected. The data is presented as both an image plot, and as a wire-frame representation. For comparison, Fig. 7(a) shows the pressure field calculated for a square plane piston using an existing method,[7] which can be compared to that shown in Fig. 6(b). By examining the normalized peak-to-peak pressure for the two methods, on a point-by-point basis, the largest discrepancy is $1.4\times10^{-8}\%$. In calculating the impulse response for a point at a give moment in time geometrically the two approaches are the same, however the way the various angles used are calculated differs and it is here that the very small discrepancy is believed to occur due to the accuracy of the algorithms used for trigonometric functions within the programming language used (Matlab). Hence, the model developed gives essentially the same results for the square source as the method already in the literature.[7] The computational time for the two ap-

TABLE III. Variables describing the irregular polygon of Fig. 10.

	Edge											
	A	B	C	D	E	F	G	H	I	J	K	L
k (mm)	2	1	0	0	−1	−2	−1	−2	2	3	2	3
l (mm)	4	2	1	1	2	2	4	3	1	4	6	4
m (mm)	1	4	0.5	−0.5	−1	0	0	−0.5	−2.5	0	0	0
θ^0	0	90	180	270	180	90	0	270	0	90	180	270

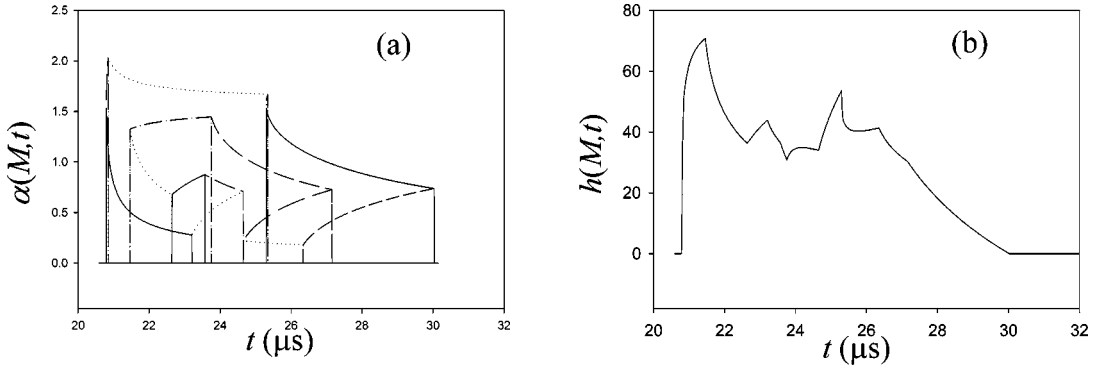


FIG. 11. The angles required to calculate the impulse response of (a) the irregular polygon using the present method, and (b) the resulting impulse response, at a location $x=2.5$ mm, $y=3$ mm, and $z=7$ mm.

proaches, using the same computer and programs written in a similar manner, was approximately the same. In addition, Fig. 7(b) shows the pressure field for a circle found using the method described by Stepanishen,[4] which is compared to the present predictions shown earlier in Fig. 6(c). It can be seen that as the number of edges contained within the polygon increases, the resultant field tends to that of the circle of Fig. 7(b), as would be expected.

Figure 8 shows predicted fields for a single field point located at $x=2.5$ mm, $y=3$ mm, and $z=7$ mm, which is outside the area of the pistons shown. For 4- and 12-sided polygons, Figs. 8(a) and (b), respectively, the impulse response $[h(M,t)]$ shows discontinuities that are related to the number of sides of the polygon. For comparison, Figs. 8(c) and (d) show the impulse response calculated for a square and circle, using existing methods[4,7] for the same field point. As seen above, the two results for the square piston are identical, whereas the result for the polygon approaches that for the circle, but with discontinuities due to the edges present.

Radiated fields for sources in the shape of (a) a triangle, (b) a square and (c) a pentagon are plotted in Fig. 9. The planes selected have been chosen such that they pass through a corner as shown. It can be seen from the resulting pressure fields that the result is symmetrical for the square source, but not for the triangle and pentagon, as expected.

IV. APPLICATION TO AN IRREGULAR POLYGON

To further demonstrate the flexibility of the technique, the concave polygon shown in Fig. 10 has been investigated (it is concave as it contains at least one internal angle of more than 180°). This has been chosen as it can be made of a series of rectangles, and as such a comparison can be made with existing techniques. The shape resembles a series of squares strung together to form a slight spiral effect. The z axis is normal to the plane of the figure. The edges of the polygon have been labeled with capital letters, and the two ends of each edge have been labeled Pt 1 and Pt 2, in the manner described in Sec. II C. These labels are shown with the use of a superscript to denote the edge letter. The origin is located at the meeting point of edge C and D. The variables required by the model to describe the polygon, namely values for k, l, m, and θ for each side, have the values given in Table III. For comparison, this transducer shape has also been modeled from four separate rectangular radiators, where the contributions from three smaller ones were subtracted from a larger one with the outer dimensions of the source (4×6 mm^2). Again the simulation was performed using a 500-kHz, 20-cycle toneburst as the piston velocity, $v(t)$, and using a speed of sound corresponding to that of air, $c=343$ ms^{-1}.

Figure 11(a) shows the values of the angles calculated for each edge using the method presented here, and Fig. 11(b) the resultant impulse response $h(M,t)$ for the radiator of Fig. 10, for the field point located at $x=2.5$ mm, $y=3$ mm, and $z=7$ mm. Figure 12(a) shows the impulse response for each of the four rectangles and Fig. 12(b) the result for the whole piston using existing methods. It can be seen by comparison of Figs. 11(b) and 12(b) that the two

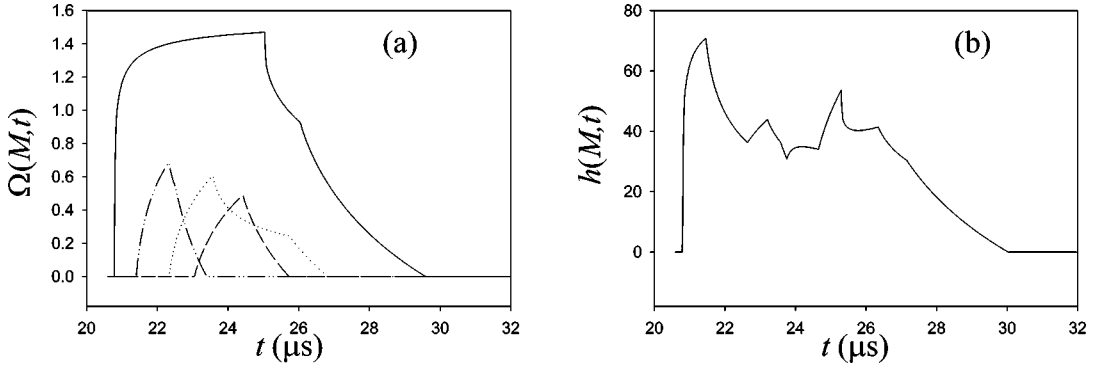


FIG. 12. The impulse response of (a) four rectangles which make up the irregular polygon using existing methods, and (b) the resulting impulse response, at a location $x=2.5$ mm, $y=3$ mm, and $z=7$ mm.

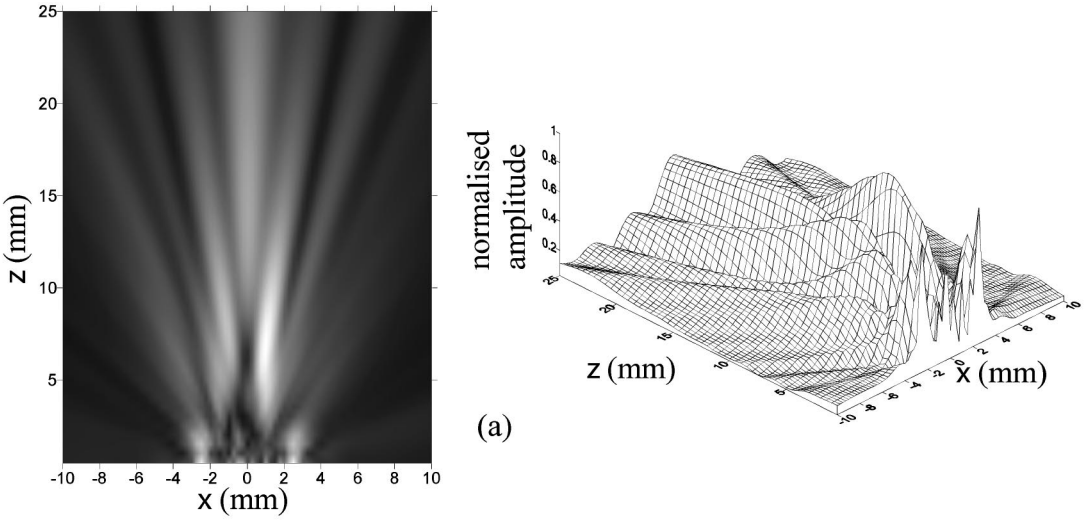


FIG. 13. The pressure field in the *xz* plane for the irregular polygon using the present method.

approaches give very similar results (differing by 2.7×10^{-6}% for a comparison of each time point). The radiated peak-to-peak pressure field has been calculated in the *xz* plane using the present method, and is shown in Fig. 13. The radiated field is complicated, as might be expected from the shape with a lack of symmetry about the *z*-axis, with major side-lobes appearing in the radiated field.

The method presented here for the irregular polygon offers a 25% computational time saving; however, it should be noted that the irregular shape used here was chosen precisely because it could be simply constructed from four rectangles and so the time saving would be expected to be much greater for other shapes. To demonstrate this consider an *n*-sided object; this can be constructed by drawing a triangle, and then from one of the sides of the triangle a further two lines forming a point; this process of drawing pairs of lines can be continued until the right number of external sides are obtained, by when in the general case $n-2$ triangles will have been drawn. Assuming that no combination of the triangles can be replaced by a rectangle, then to find the impulse response at a single field point a calculation must be performed for n-2 triangles (using the method described by Jensen[10]) or *n* sides. For each of the triangles which form a single external edge, the time period over which the impulse response must be considered is greater than for that external edge; for the two triangles with two external edges, the time period is greater than for each of the edges. In addition the calculation of the impulse response for each triangle is considerably more complicated than that for each edge. Bear in mind that the time taken for the calculation of the field from a square directly was comparable to that using four edges; had that four-sided shape been irregular, then two triangles and so two direct calculations would be needed rather than four edges using the method presented here.

V. CONCLUSIONS

A model has been presented which is capable of calculating the impulse response at any field point from a planar piston of polygonal shape. By convolution with the time derivative of the piston velocity and multiplication with the fluid density, this impulse response can be used to calculate the resultant pressure field. The modeling method has been described for an arbitrary polygon and demonstrated for a variety of shapes.

For the purpose of validation, pressure fields and the impulse response for individual field points have been compared and shown to be in agreement with existing methods. This has been done for a square piston, and a concave polygon, which can be made up from four rectangles. In addition, the field from a 12-sided regular polygon was shown to approximate that from a circular plane piston. A further analysis was presented, which demonstrated that the technique could be used for an irregular polygon. This showed that the technique could be used for a wide range of acoustic source shapes, provided that they can be broken down into a series of straight edges. The flexibility and simplicity of implementation of this model will make it of use to groups working in the area of pressure field modeling and transducer design.

[1] J. W. S. Rayleigh, *The Theory of Sound*, Vol. 2 (Dover Publications, New York, 1945), p. 107.
[2] G. R. Harris, "Review of transient field-theory for a baffled planar piston," J. Acoust. Soc. Am. **70**, 10–20 (1981).
[3] F. Oberhettinger, "On transient solutions of the baffled piston problem," J. Res. Natl. Bur. Stand. **65**, 1–6 (1961).
[4] P. R. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," J. Acoust. Soc. Am. **49**, 1629–1638 (1971).
[5] A. Freedman, "Sound field of a rectangular piston," J. Acoust. Soc. Am. **32**, 197–209 (1960).
[6] J. C. Lockwood and J. G. Willette, "High-speed method for computing the exact solution for the pressure variations in the nearfield of a baffled piston," J. Acoust. Soc. Am. **53**, 735–741 (1973).
[7] J. L. San Emeterio and L. G. Ullate, "Diffraction impulse-response of rectangular transducers," J. Acoust. Soc. Am. **92**, 651–662 (1992).
[8] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 262–267 (1992).
[9] P. Faure, D. Cathignol, and J. Y. Chapelon, "Diffraction impulse response of arbitrary polygonal plane," Acta Acust. **2**, 257–263 (1994).
[10] J. A. Jensen, "Ultrasound fields from triangular apertures," J. Acoust. Soc. Am. **100**, 2049–2056 (1996).
[11] D. A. Hutchins, H. D. Mair, P. A. Puhach, and A. J. Osei, "Continuous-wave pressure fields of ultrasonic transducers," J. Acoust. Soc. Am. **80**, 1–11 (1986).

[12] R. Eriksson, A. Holm, M. Landeborg, H. W. Persson, and K. Lindström, ''The 10 MHz ultrasonic near-field: calculations, hydrophone and optical diffraction tomography measurements,'' Ultrasonics **31**, 439–446 (1993).
[13] T. J. Robertson, A. Neild, D. A. Hutchins, J. S. McIntosh, D. R. Billson, R. A. Noble, R. R. Davies, and L. Koker, ''Radiated fields of rectangular air-coupled micromachined transducers,'' in *Proceedings 2001 IEEE International Ultrasonics Symposium*, Atlanta, GA, Vol. 2, pp. 891–894.
[14] A. Neild, D. A. Hutchins, T. J. Robertson, L. A. J. Davis, and D. R. Billson, ''The radiated fields of ultrasonic phased arrays,'' submitted to Ultrasonics.
[15] M. Akhnak, O. Martinez, L. G. Ullate, and F. Montero de Espinosa, ''64 Element two-dimensional piezoelectric array for 3D imaging,'' Ultrasonics **40**, 139–143 (2002).
[16] R. Reibold and R. Kazys, ''Radiation of a rectangular strip-like focussing transducer, Part 1: harmonic excitation,'' Ultrasonics **30**, 49–55 (1992).
[17] T. J. Robertson, D. A. Hutchins, and D. R. Billson, ''An air-coupled line-focussed capacitive ultrasonic transducer,'' in *Proceedings 2000 IEEE International Ultrasonics Symposium*, San Juan, Puerto Rico, Vol. 2, pp. 1061–1064.
[18] E. Skudrzyk, *The Foundations of Acoustics* (Springer-Verlag, Wien, 1971), pp. 512–556.
[19] A. Neild and D. A. Hutchins, ''A theoretical model for the effect of a finite-sized acoustic receiver,'' (accepted) J. Acoust. Soc. Am.

Ultrasonic delineation of aortic microstructure: The relative contribution of elastin and collagen to aortic elasticity

Jon N. Marsh
Washington University School of Medicine, Cardiovascular Division, St. Louis, Missouri 63110

Shin Takiuchi
National Cardiovascular Center, 5-7-1 Fujishirodai Suita, Osaka, 565-0873, Japan

Shiow Jiuan Lin, Gregory M. Lanza, and Samuel A. Wickline
Washington University School of Medicine, Cardiovascular Division, St. Louis, Missouri 63110

(Received 20 March 2003; revised 7 February 2004; accepted 10 February 2004)

Aortic elasticity is an important factor in hemodynamic health, and compromised aortic compliance affects not only arterial dynamics but also myocardial function. A variety of pathologic processes (e.g., diabetes, Marfan's syndrome, hypertension) can affect aortic elasticity by altering the microstructure and composition of the elastin and collagen fiber networks within the tunica media. Ultrasound tissue characterization techniques can be used to obtain direct measurements of the stiffness coefficients of aorta by measurement of the speed of sound in specific directions. In this study we sought to define the contributions of elastin and collagen to the mechanical properties of aortic media by measuring the magnitude and directional dependence of the speed of sound before and after selective isolation of either the collagen or elastin fiber matrix. Formalin-fixed porcine aortas were sectioned for insonification in the circumferential, longitudinal, or radial direction and examined using high-frequency (50 MHz) ultrasound microscopy. Isolation of the collagen or elastin fiber matrices was accomplished through treatment with NaOH or formic acid, respectively. The results suggest that elastin is the primary contributor to aortic medial stiffness in the unloaded state, and that there is relatively little anisotropy in the speed of sound or stiffness in the aortic wall.
 [DOI: 10.1121/1.1698887]

PACS numbers: 43.35.Cg, 43.58.Dj, 43.80.Ev [FD] Pages: 2032–2040

I. INTRODUCTION

Aortic elasticity is an important factor in maintaining circulatory function. The inherent elastic behavior of the aorta arises primarily from the composition and structure of the tunica media, which is composed of concentric layers of elastic lamellar units.[1] These lamellae consist of smooth muscle cells, elastin fibers, and collagen fibrils. Elastin is a remarkably extensible fibrous protein that is a primary determinant of aortic distensibility,[2] while the mesh of collagen fibrils provides ultimate protection against overextension and disruption of the vessel under conditions of excessive strain.[1]

A variety of factors that elicit functional changes in elastin and collagen contribute to compromised aortic elastic behavior. Both the amount of connective tissue and the ratio of collagen to elastin increase in the aorta as a result of the aging process.[2] Increased stiffness in the aorta is also found in patients with diabetes mellitus independent of other vascular complications,[3] and is potentially a result of the accumulation of ''advanced glycosylation end-products'' (AGEs) that promote abnormally high levels of collagen cross-linking.[4,5] Additionally, connective tissue diseases such as Marfan's syndrome exhibit decreased elastin content and abnormalities in elastin structure and organization, which can cause aortic dissection through increased tissue fragility.[6]

In view of the role the aorta plays in modulating circulatory flow and ventricular loading, aortic elasticity may represent an important diagnostic metric for defining overall hemodynamic health.[5] A variety of techniques and modalities have been utilized to quantify aortic elastic properties. While measures of elasticity can be acquired *in vitro* by mechanical techniques,[2,7] quantification *in vivo* is typically obtained indirectly either by analysis of the pulse wave velocity in the aorta[8–10] or simultaneous measurement of aortic diameter and internal pressure during the heart cycle.[11–15] Such measurements may be either difficult to localize to a specific portion of the aorta or to obtain routinely *in vivo*.

As a tissue characterization modality, ultrasound potentially could offer a direct, nondestructive measurement of the mechanical properties of soft tissue. In this study we address two aspects of ultrasonic tissue characterization pertinent to assessment of vascular material properties: (1) the contributions of elastin and collagen to its elasticity, and (2) the extent to which its anisotropy is manifest in measurements of speed of sound and stiffness. The velocity at which longitudinal sound waves propagate in a specific direction through a linear elastic medium is related to its elasticity or ''stiffness'' in that direction through the general relation $c=\rho v^2$, where c is the elastic constant (also called the elastic stiffness coefficient), ρ is the density, and v is the acoustic phase velocity. We sought to define the contribution of selected components of the arterial wall (collagen and elastin) to its mechanical properties by measuring the magnitude and directional dependence (anisotropy) of the speed of sound and elasticity before and after chemical treatments to remove either collagen or elastin components, using techniques demonstrated in previous work from our group.[16] In this manner,

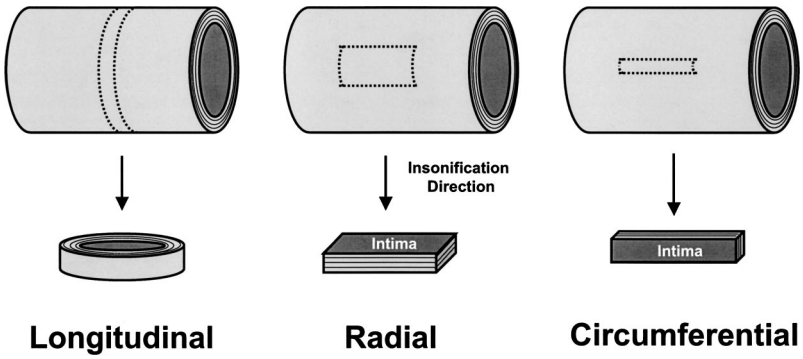


FIG. 1. Illustration of orientation of aortic sections used for acoustic microscopy. The upper part of the figure depicts cylindrical sections of excised aorta, and dotted lines indicate the portions of the aortic wall that were sliced out for acoustic microscopy. Cross-sectional slices were labeled ''Longitudinal'' and were insonified parallel to the vessel axis. ''Radial'' samples were sliced from a portion of the vessel wall, and the direction of insonification was from the intimal side to the adventitia. ''Circumferential'' samples were also sliced from the wall, but were oriented to expose the medial fiber matrix to insonification in a tangential manner. Arrows indicate the direction of insonification in each case.

the three-dimensional elastic behavior of the collagen-isolated and elastin-isolated matrices could be compared with that of the intact, unmodified aortic tissue.

II. MATERIALS AND METHODS

A. Sample preparation

Fresh descending thoracic aortas were obtained from male domestic pigs, cleaned of extraneous connective tissue, and immediately immersed in 10% buffered formalin. After fixation was complete, the aortas were prepared for scanning with acoustic microscopy. The samples were sectioned in one of three ways, according to the desired orientation of the sample with respect to the insonification direction (see Figs. 1 and 2). The first group of samples was used for longitudinal insonification by making parallel cuts through the aortas perpendicular to the vessel axis to produce 2-mm-thick circular sections. The circular sections were mounted flat on a polished steel plate for insonification of the aortic tissue along the vessel axis.

The second group of samples was used for radial insonification by cutting rectangular (approximately 2×5 mm^2) sections out of the aortic wall. These sections were mounted flat with the intimal side up for insonification. It should be noted that in this orientation, the acoustic pathway through the sample included the intima and a thin layer of adventitia. The intimal layer of healthy vasculature is exceedingly thin and not expected to contribute to the aorta's mechanical properties. Because of the geometry and small size of the samples, it was difficult to mechanically remove the adventitia in its entirety; however, the bulk of the tissue in such samples was from the medial layer, and thus deemed to yield an accurate representation of medial acoustic properties.

The last group of samples, used for circumferential insonification, were obtained by cutting out a 2-mm-wide strip of the aortic wall. These samples were mounted such that the intimal and adventitial surfaces were exposed on the sides, while the tangential cross-sectional surfaces were situated on the top and bottom. In total, there were 14 circumferential specimens, 12 longitudinal specimens, and 13 radial specimens used for baseline ultrasound measurements.

After baseline measurements were obtained, samples from each group were subjected to 7 days treatment with either 10% NaOH or 88% formic acid to isolate either the collagenous[17] or elastin[18] extracellular components, respectively. Once the tissue digestion treatment was complete, the samples were rinsed and allowed to equilibrate in distilled water for 24 h before reinsonification.

B. Velocity measurements

Ultrasound measurements were performed with a custom-designed acoustic microscope. A high-frequency pulser/receiver unit (Panametrics 5900, Waltham, MA) was used to excite a 50-MHz, 0.5-in. focus, 0.25-in.-diam transducer (Panametrics V390), which produced short-duration pulses for pulse-echo tissue insonification. The transducer's beam diameter (nominally 60 μm at the -6 dB level) permitted the acquisition of several hundred independent measurements for each sample. The received signals were amplified (Panametrics 5900) and digitized to 8 bits (HP 54510B, Hewlett-Packard, Palo Alto, CA) at 500 megasamples/s.

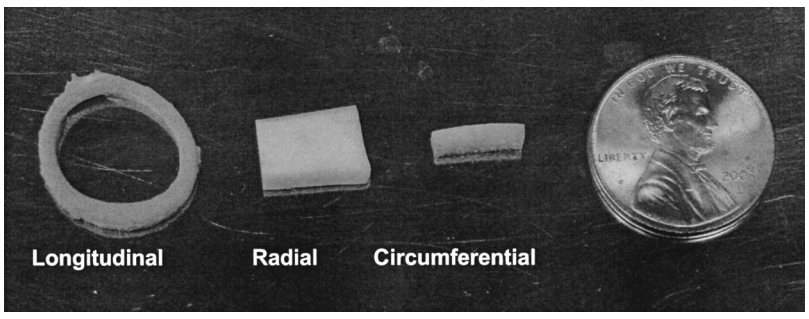


FIG. 2. Photograph of typical porcine aortic sections used for insonification in longitudinal, radial, and circumferential directions. Insonification direction is into the figure.

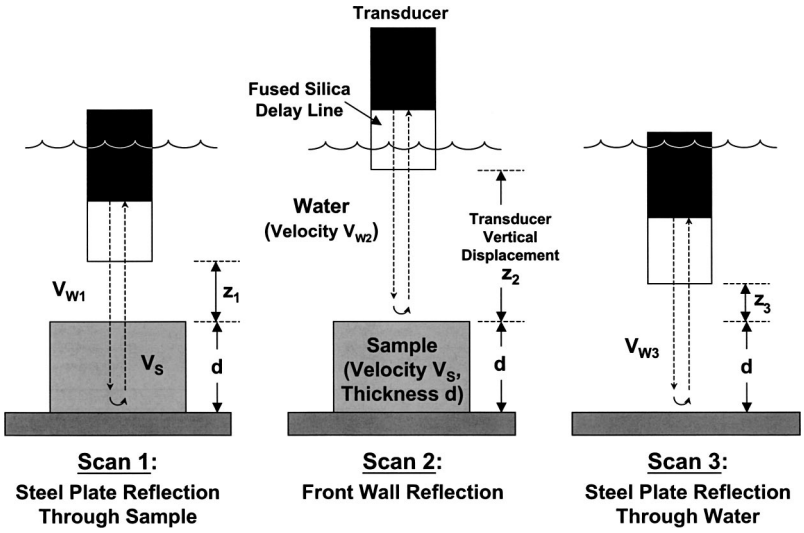


FIG. 3. Schematic representation of transducer and specimen configuration for each of the three scans used to determine speed of sound within the sample. In ''scan 1,'' transducer is focused on steel reflector with sample intervening. In ''scan 2,'' transducer is moved so that focus is near surface of the sample. Finally, the sample is removed and the transducer focused again on steel plate in ''scan 3.''

Each 1024-point waveform was averaged 128 times in the time domain and then stored on a computer (Macintosh G4) for offline analysis. Each sample was scanned in a two-dimensional grid (typically 300 points, depending on sample size) with a step size of 100 μm. Scan areas for the longitudinal and circumferential samples were chosen to include only regions within the tunica media.

Three scans were performed for each specimen to determine the thickness d and group velocity v_s at each point within the sample (Fig. 3). In the first scan (''scan 1''), the sample was mounted on the steel plate reflector and waveforms were acquired with the transducer positioned vertically such that its focus was at the level of the steel plate. The next scan (''scan 2'') traversed the same points on the sample, but the transducer was moved so that its focus was at the surface of the tissue, and the front wall reflection was acquired at each point. The same coordinates were again used for the final scan (''scan 3''), but in this case the sample was removed, and the transducer was focused on the steel plate. The position of the scanning gantry's Z-axis (z_1, z_2, or z_3 in Fig. 3) was recorded for each scan to determine the relative vertical displacement of the transducer between scans. Water temperature within the tank was also recorded at the beginning of each scan in order to determine the speed of sound in the host medium (v_{w1}, v_{w2}, or v_{w3}), according to a polynomial fit for the speed of sound in pure water.[19] All scans were performed at room temperature, which ranged from 18.5 to 22.5 °C.

A custom program written in LabVIEW (National Instruments, Austin, TX) was used to automatically determine the time-of-flight of the returned echoes for each scan. The procedure is illustrated for example waveforms in Fig. 4. First, the envelope of each waveform (defined by the magnitude of the analytic signal) was calculated. A peak-detection algorithm was applied to the envelope to determine the height and width of the first peak detected above a minimum amplitude threshold. The echo arrival time was defined as the temporal position of the first point of the waveform envelope to attain a value equal to or greater than half the first peak's magnitude. This procedure was used to determine the round-trip time of flight (t_1, t_2, and t_3) for the acoustic pulse at each point in the sample in each of the three scans. The total time of flight is related to the distance of propagation in scans 1, 2, and 3 by the following formulas:

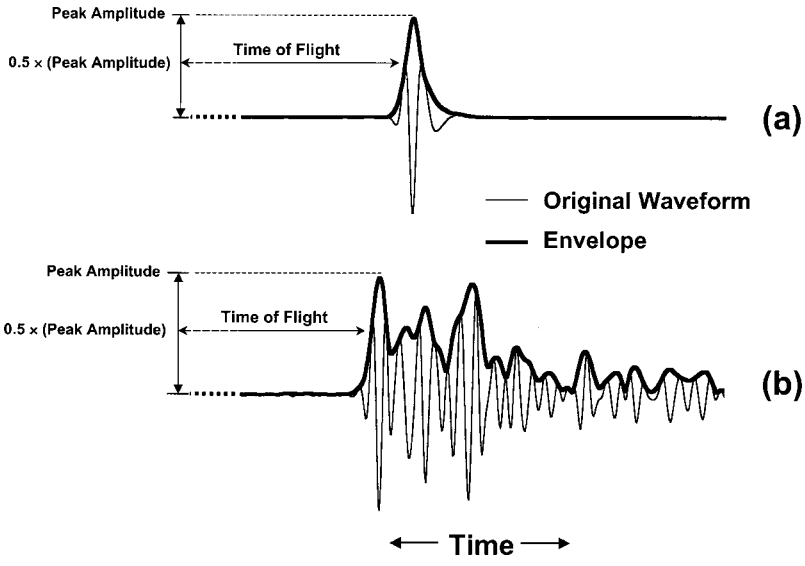


FIG. 4. Example waveforms used to determine echo arrival times for sample thickness and velocity calculations. Unprocessed radiofrequency data are shown with thin lines, and corresponding envelopes (i.e., magnitude of the analytic signal) are displayed with bold lines. (a) Representative waveform and envelope from ''scan 1''; (b) representative waveform and envelope from ''scan 2.'' Vertical scales of (a) and (b) are different. Echo arrival time was determined in each case by finding the temporal position of the first point of the waveform envelope to attain a value equal to or greater than half the first peak's magnitude.

$$t_1 = 2\left(\frac{z_1}{v_{w1}} + \frac{d}{v_s}\right) + t_{\text{delay}}, \tag{1}$$

$$t_2 = 2\frac{z_2}{v_{w2}} + t_{\text{delay}}, \tag{2}$$

$$t_3 = 2\left(\frac{z_3 + d}{v_{w3}}\right) + t_{\text{delay}}, \tag{3}$$

where t_{delay} is the transducer delay line round-trip propagation time. Upon rearranging variables and then subtracting Eq. (2) from Eq. (3), the tissue sample thickness d at each point in the sample can be calculated in terms of known values for the speed of sound in water, the measured times-of-flight, and the difference in vertical displacement:

$$d = \tfrac{1}{2}v_{w3}(t_3 - t_{\text{delay}}) - \tfrac{1}{2}v_{w2}(t_2 - t_{\text{delay}}) + (z_2 - z_3). \tag{4}$$

Similar manipulation [i.e., rearranging variables, subtracting Eq. (2) from Eq. (1), and substituting of Eq. (4) for d] then yields the speed of sound v_s at each position in the sample:

$$v_s = v_{w1}\left[\frac{\frac{1}{2}v_{w3}(t_3 - t_{\text{delay}}) - \frac{1}{2}v_{w2}(t_2 - t_{\text{delay}}) + (z_2 - z_3)}{\frac{1}{2}v_{w1}(t_1 - t_{\text{delay}}) - \frac{1}{2}v_{w2}(t_2 - t_{\text{delay}}) + (z_2 - z_1)}\right]. \tag{5}$$

The speed of sound for each sample was defined as the mean value of v_s averaged over every site scanned in the sample.

C. Density measurements

Density measurements were performed on formalin-fixed sections of descending thoracic porcine aorta (approximately 5 cm in length). A balance (Mettler Toledo PB303, Columbus, OH) with specific gravity kit (Sartorius YDK 01, Goettingen, Germany) was used to determine the density first on the intact specimens ($n=10$), and then again after one week of digestion in either NaOH ($n=5$) or formic acid ($n=5$). In the case of the intact and elastin-isolated specimens, the samples were immersed in distilled water for at least an hour prior to measurement. Each sample was then removed from the water and patted dry to remove surface moisture, weighed in air, then weighed while submerged in a liquid of known density (distilled water) to determine its specific gravity. This technique was modified for the collagen-isolated samples because their relatively similar density value as compared to water and somewhat gelatinous consistency (as a result of the digestion treatment) made it difficult to obtain consistent readings either in water or air. For these cases, sample weights were determined in both *n*-octane and toluene, whose lower densities (0.703 and 0.865 g/mL at 22 °C, respectively) permitted more robust measurement of the submerged sample weights.

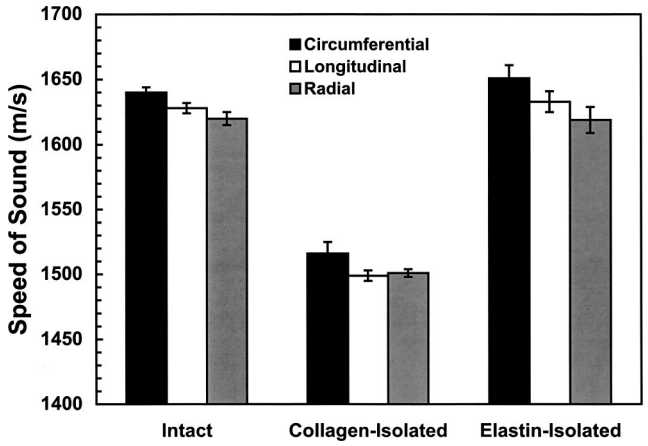


FIG. 5. Speed of sound (m/s) in fixed porcine aortic media as a function of direction of insonification and treatment. Error bars represent standard errors for the measurements.

D. Statistical analysis

Paired *t*-tests were used to establish the statistical significance ($p<0.05$) of the null hypothesis for differences in results for density and speed of sound between samples grouped according treatment type and orientation. Analysis of variance (ANOVA) was performed using Scheffe's multiple comparison procedure on the stiffness coefficient values to test for significant anisotropy (at a significance level of 0.05) within each treatment group, as well as between treatment groups for similar directions of insonification.

III. RESULTS

A. Density

Densities for intact, elastin-isolated, and collagen-isolated tissues were 1.072 ± 0.002 g/cm^3 ($n=10$), 1.098 ± 0.004 g/cm^3 ($n=5$), and 1.04 ± 0.02 g/cm^3 ($n=5$), respectively. Statistically significant differences existed between the elastin-isolated group and either the intact or collagen-isolated groups. There was no meaningful difference between the intact and collagen-isolated groups with regard to density, however.

B. Speed of sound

Results for speed of sound for all tissue orientations and treatments are shown graphically in Fig. 5. Error bars repre-

TABLE I. Results of paired *t*-tests for significant differences between speed of sound measured for sample groups insonified in different tissue orientations. Asterisks ("*") denote significant differences at the $p<0.05$ level; "ns" denotes no significant difference between groups. Sample pairs are grouped according to tissue treatment.

Orientation comparison	Intact	Collagen-isolated	Elastin-isolated
Circumferential versus Longitudinal	ns	ns	ns
Longitudinal versus Radial	ns	ns	ns
Radial versus Circumferential	*	ns	ns

TABLE II. Results of paired *t*-tests for significant differences between speed of sound measured for sample groups subjected to different digestion treatments. Asterisks (''*'') denote significant differences at the $p < 0.05$ level; ''ns'' denotes no significant difference between groups. Sample pairs are grouped according to direction of insonification.

Treatment comparison	Circumferential	Longitudinal	Radial
Intact versus Collagen-Isolated	*	*	*
Intact versus Elastin-Isolated	ns	ns	ns
Collagen versus Elastin-Isolated	*	*	*

sent standard errors for the measurements. Statistical differences between pairs of sample groups are denoted by asterisks in Tables I and II.

1. Anisotropy

There is no significant difference in speed of sound in any of the three orientations for either of the treated sample groups; the same holds true for the intact tissue, with the exception of the radial and circumferential sound speeds, which differentiate to a modest degree.

2. Contributions from elastin and collagen

No statistically significant difference exists between the intact and elastin-isolated tissue for any orientation. However, the collagen-isolated samples exhibit substantially lower group velocity than did their intact or elastin-isolated counterparts. In fact, the speed of sound for the collagen-isolated samples is very nearly equal to that of distilled water at room temperature.

C. Stiffness coefficients

Figure 6 illustrates computed values for the stiffness coefficients of intact, elastin-isolated, and collagen-isolated aorta measured in the circumferential, longitudinal, and radial directions. The values were computed as the product $\overline{\rho} \cdot \overline{v}^2$, where $\overline{\rho}$ is the mean density of the appropriately treated aortic segments given above, and $\overline{v}$ is the mean speed of sound of each group of samples. Error bars represent propagated errors from the density and velocity measurements.

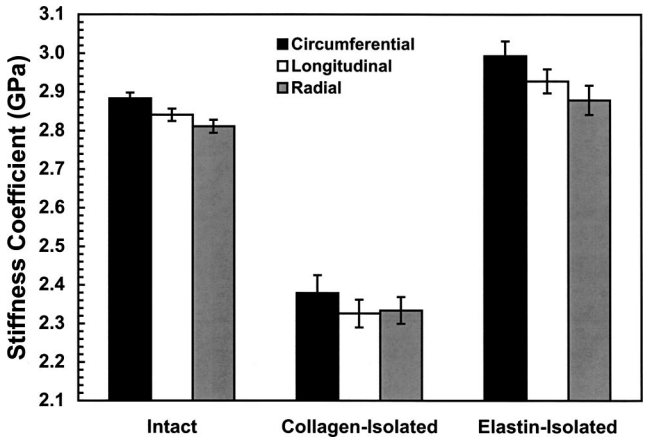


FIG. 6. Stiffness coefficients (GPa) of fixed porcine aortic media determined for specific orientations and treatments. Error bars represent propagated errors from density and speed of sound measurements.

1. Anisotropy

There was no significant directional dependence for stiffness in either the collagen- or elastin-isolated samples; the only significant difference in the intact group existed between circumferential and radial orientations.

2. Contributions from elastin and collagen

As shown in Tables III and IV, isolation of either collagen or elastin produced a statistically significant difference in the stiffness coefficient for all orientations except radial, in which there was no significant difference between intact and elastin-isolated tissue.

IV. DISCUSSION

The complex structure of the aorta enables it to function as more than a simple conduit for blood distribution to the body, but also to modulate the dynamics of blood flow to increase the effectiveness of the pumping heart. Normally, the aorta distends rapidly during ventricular ejection to transiently accommodate more than 50% of the cardiac stroke volume.[2] Because of its intrinsic elasticity, most of the potential energy stored in the stretched aorta is returned during diastole, which provides an additional passive pumping mechanism while the heart is filling. This elastic behavior serves to ''smooth'' the initial pulsatile flow and thus maintain a relatively constant blood pressure downstream of the heart, and also modulates the hydraulic load experienced by the left ventricle by effectively matching its output to the dynamic flow impedance of the arterial tree over the duration of the heart cycle.[20]

Previous ultrasonic studies of aortic tissue have focused on the attenuation and backscatter to characterize its physical state. These indices can differentiate normal aortic tissue segments from regions exhibiting various stages of atherosclerotic plaque formation.[21–27] Similarly, the abnormal aortic architecture associated with Marfan's syndrome manifests decreased ultrasonic backscatter relative to aortas with normal elastic tissue structure.[28] Ultrasound tissue characterization methods also register the anisotropic geometry of elastin and collagen fiber lattices in aorta and other elastic arteries, which exhibit variations in backscatter and attenuation that depend on the direction of insonification,[29–31] and that relate to microscopic fiber organization. For example, Hall *et al.* demonstrated that elastin-isolated canine aortas exhibit a backscatter coefficient several times larger than do collagen-isolated samples when insonified in the radial direction.[16]

In the present study, we sought to extend these observations by defining the relative contributions of collagen and

TABLE III. Results of analysis of variance (ANOVA) performed using Scheffe's multiple comparison procedure to test for significant differences between stiffness coefficients calculated for sample groups insonified in different tissue orientations. Asterisks (''*'') denote significant differences at the $p<0.05$ level; ''ns'' denotes no significant difference between groups. Sample pairs are grouped according to tissue treatment.

Orientation comparison	Intact	Collagen-Isolated	Elastin-Isolated
Circumferential versus Longitudinal	ns	ns	ns
Longitudinal versus Radial	ns	ns	ns
Radial versus Circumferential	*	ns	ns

elastin fibers to the elastic behavior of aorta in its unloaded state with computations of fundamental elastic stiffness moduli. The elastic stiffness coefficients, also called elastic constants, are defined as the constants of proportionality between the stress and strain matrices in the generalized form of Hooke's law for linear elastic solids.[32] The form of the generic equation $c=\rho v^2$ is derived from this relationship in combination with Newton's second law.[32] The stiffness coefficients in general form a fourth-rank tensor with 21 independent components for an anisotropic medium, although materials with certain symmetries may have far fewer (e.g., isotropic, homogeneous materials are fully described with only two independent elastic stiffness coefficients). Complete sets of measurements of these elastic constants may be used to determine other derivative indices, such as Young's modulus. Toward this end, we subjected samples of porcine aorta to treatment with either NaOH or formic acid to isolate the collagen and elastin extracellular matrices, respectively, and measured the speed of sound of compression waves in three orthogonal directions to characterize anisotropy of elastic stiffness coefficients of these components.

The most striking result is the significantly lower speed of sound, and hence reduced elastic stiffness modulus, of the collagen-isolated tissue relative to the intact or elastin-isolated samples. This lower stiffness occurs irrespective of insonification direction, whereas the elastin-isolated tissues exhibit somewhat comparable stiffness coefficients to the intact samples. It is interesting to note that this occurs despite the fact that the elastic modulus for isolated collagen fibers is many times greater than for elastin.[2] This result can be interpreted according to the expected fiber arrangement in unloaded elastic arteries. In this ''low-strain'' regime, elastic behavior is governed primarily by the elastin network, and the collagen fiber mesh exists in its ''crimped'' or unstretched state and contributes little to the overall stiffness.[7,33,34] As the deformations induced by ultrasound are extremely small, and the aortic tissue was not exposed to external stresses, the low-strain limit applies to the measurements shown here. Thus, elastin dominates the material properties of the aorta in this state.

It is interesting to note that there is much less apparent anisotropy in the stiffness coefficients than is exhibited by other ultrasound tissue characterization indices for aorta. For example, the results given by Nguyen *et al.*[31] indicate that the attenuation coefficient is greater in the circumferential and longitudinal orientations relative to radial measurements. Likewise, the backscatter coefficient is greater in the radial direction than in either longitudinal or circumferential directions by approximately two orders of magnitude. The organization of elastin into numerous concentric elastic lamellae has been postulated to account for the large anisotropy in backscatter coefficient.[31] These fibrous lamellae, typically 2 to 3 μm thick,[18] are separated by approximately 15 μm[35] and exhibit fenestrations through which less densely packed tangles of elastin and collagen fibers pass to tie the structure together radially.[7] While high-frequency ultrasonic backscatter is sensitive to the local variations in acoustic impedance that arise from the lamellae and intervening gaps,[36] the group velocity across the entire sample is predominantly a function of the *net* contributions of the variously oriented microscopic components of the elastic fiber network acting to propagate the acoustic wave through the tissue. As compared with the more prominent anisotropy of backscatter and attenuation, the very moderate directional dependence of the stiffness coefficient values suggests that, in the unloaded state, the relative population density of elastin fibers aligned in the circumferential, longitudinal, and radial directions may be similar throughout the bulk of the media.

In order to verify whether the measurement techniques utilized in this study are in fact sensitive enough to resolve small differences in velocities (and, hence, in stiffness coefficients), we must examine the sensitivity and uncertainty of the velocity measurements. Based on Eq. (1), one can estimate the difference in acoustic transit time t_1 resulting from a change in propagation speed through the tissue. The partial derivative of Eq. (1) with respect to the speed of sound in the

TABLE IV. Results of analysis of variance (ANOVA) performed using Scheffe's multiple comparison procedure to test for significant differences between stiffness coefficients calculated for sample groups subjected to different digestion treatments. Asterisks (''*'') denote significant differences at the $p<0.05$ level; ''ns'' denotes no significant difference between groups. Sample pairs are grouped according to direction of insonification.

Treatment comparison	Circumferential	Longitudinal	Radial
Intact versus Collagen-isolated	*	*	*
Intact versus Elastin-isolated	*	*	ns
Collagen-versus Elastin-isolated	*	*	*

sample, v_s, is given by $\partial t_1/\partial v_s = -2d/v_s^2$; for typical values of thickness d (2 mm) and v_s (1600 m/s), this yields a change of 1.6 ns in transit time for a change of 1 m/s in speed of sound. This is approximately equal to the sampling interval for these measurements (2 ns), which is the minimum detectable time shift. Time shifts this small were detectable in the waveform envelopes because of the narrow character of the echo peaks (<40 ns, half-width at half-max), as well as the insignificant amount of observable jitter associated with the acquisition and digitizing process. Although the timing measurements are potentially sensitive enough to indicate velocity differences as small as 1 to 2 m/s, a truer measure of the sensitivity of the technique can be obtained by considering the uncertainty in the measurements that are input to the speed of sound, as given in Eq. (5). As stated above, we estimate the uncertainty of the threshold-detected transit time for the reflected echoes to be equal to the digitizer's sampling interval, which is 2 ns. The scanning gantry's stepper motors, used to move the precision linear stages to which the transducer was affixed, employ encoders and feedback control to accurately register and maintain the stages' positions when at rest. These stages have a nominal repeatability of 1 μm; however, we conservatively estimate the uncertainty in the transducer's vertical position to be ± 5 μm. The final quantity to be estimated is the temperature of the waterbath (used to calculate the speed of sound in water through the polynomial expression given in the reference by Del Grosso and Mader[19]); this uncertainty is taken to be 0.1 °C. Propagation of these uncertainties based on Eq. (5) typically yields a resultant uncertainty of about 8 m/s in the speed of sound, which is roughly equal in magnitude to the standard errors of the velocity results. Thus, the results may be considered reliable to within the bounds of the error bars.

A variety of mechanical indices have been used to assess the elasticity of aortic tissue, including so-called ''pressure-strain'' and ''circumferential'' elastic moduli, which can be applied to *in vivo* measurements of stress and strain (i.e., pulse pressure and changes in vessel diameter). Other traditional engineering measures such as Young's modulus can be derived from each these.[1] In this study we have computed specific elastic stiffness coefficients from measurements of speed of sound and mass density of aortic tissue. Although the various forms of elastic moduli are expressed in the same units as stiffness coefficients, direct comparison between the two requires further information. In the case of isotropic materials, for example, one must also determine the Poisson's ratio or, alternatively, the shear modulus (determined from the speed of shear wave propagation). For highly ordered anisotropic tissues, the number of independent parameters required to completely specify elastic behavior grows considerably. Even the relatively simple structure of tissues with a uniaxial fiber arrangement (e.g., tendon) requires measurement of five independent stiffness coefficients to completely define the linear elastic behavior, necessitating measurement of speed of sound of both compressional and transverse waves through the tissue at specific orientations.[37–39] In such cases the expression for Young's modulus along an arbitrary direction is a function of all the elastic constants. Because of the somewhat complex arrangement of circumferentially oriented collagen fibers and sheets of elastin within aorta, it may be necessary to specify a greater number of elastic coefficients to completely describe its elastic behavior. The measurements presented here for aorta insonified in three orthogonal directions yield three of these coefficients.

Direct comparisons of the data in this study can nevertheless be made with other published data in the case of the speed of sound in aorta. Geleskie and Shung[40] reported a value of 1572 m/s in fresh and frozen/thawed bovine upper descending aorta at 20 °C, although tissue orientation was not given. Lockwood *et al.*[30] reported velocities ranging from 1579 to 1628 m/s for frozen/thawed human arteries measured at 37 °C in the radial direction. Saijo *et al.*[27] published a value of 1614 m/s for frozen/thawed fixed normal human aortic media insonified longitudinally. These studies and the present one indicate that the speed of sound in elastic vessels is somewhat greater than the average value of 1540 m/s assumed for most soft tissue. In comparing these values with those of the present study, the variation in species and sample preparation should be noted. In particular, donor age could exert a significant effect, because the ratio of collagen to elastin in aortic media increases in senescence.[2]

Although tissue fixation was required for this study to ensure the samples' integrity for ultrasonic measurements performed before and after seven days of treatment with NaOH or formic acid, the effects of fixation on the elastic properties of aorta must be considered. Park *et al.*[41] found that whole, noncalcified, excised human femoral arteries became 68% less distensible when exposed to physiologic pressures after formalin fixation. However, muscular arteries such as the femoral possess a significantly different medial structure than do elastic arteries (e.g., aorta), and contain a greater proportion of collagen and smooth muscle. *In vitro* studies have shown that formalin fixation elicits a only a small change (on the order of 1% decrease) in the speed of sound of liver, spleen, and brain,[42,43] suggesting that low-strain stiffness measurements are not affected to an appreciable degree.

It is possible that differences between low- and high-strain behavior of fixed samples occurs because formalin fixation affects tissue components heterogeneously. The elastic properties of elastin are affected little, if at all, by exposure to formalin,[7] whereas the chemical changes induced by formalin cross-linking of myocardial collagen have been proposed as a model of age- and disease-related stiffening of the heart.[44] If the stiffness of collagen would be affected by fixation moreso than the stiffness of elastin, then we might anticipate that any differences in intact aorta would only be observable when the collagen fibers are recruited in the high-strain limit of vessel deformation.

These considerations in sum suggest that the elastin component of aortic media is the predominant source of interest in ultrasound velocity measurements under linear, unloaded conditions. Typically, mechanical tests performed on intact aortas demonstrate that the degree of elastic anisotropy increases as the elastic moduli become increasingly nonlinear with increasing distension.[45–48] Such behavior presumably indicates increasing collagen recruitment and mechanical function transitioning to the stiffer, high-strain regime.

V. CONCLUSIONS

We have shown that the speed of sound in aorta can be a useful tool for characterizing its stiffness coefficients. Measurements performed in orthogonal directions on intact, elastin-isolated, and collagen-isolated samples demonstrated that (a) elastin is the primary contributor to aortic medial stiffness in the unloaded state, and (b) there is little significant anisotropy in the unloaded stiffness of the medial layer. Despite the limitations of the current measurements, the differences in elastic behavior in the low-strain regime appear significantly affected by elastin as compared with collagen. Thus, as a method of direct interrogation of vascular stiffness, ultrasound tissue characterization methods would appear to offer unique insights into normal and pathological structure and function, and merit further effort to achieve clinical implementation.

ACKNOWLEDGMENTS

The authors would like to thank John S. Allen and Mark D. McLean for their surgical assistance.

[1] F. H. Silver, D. L. Christiansen, and C. M. Buntin, ''Mechanical properties of the aorta: a review,'' Crit. Rev. Biomed. Eng. **17**, 323–358 (1989).

[2] P. B. Dobrin, ''Mechanical properties of arteries,'' Physiol. Rev. **58**, 397–460 (1978).

[3] J. Hu, M. Wallensteen, and G. Gennser, ''Increased stiffness of the aorta in children and adolescents with insulin-dependent diabetes mellitus,'' Ultrasound Med. Biol. **22**, 537–543 (1996).

[4] M. Brownlee, A. Cerami, and H. Vlassara, ''Advanced glycosylation end products in tissue and the biochemical basis of diabetic complications,'' N. Engl. J. Med. **318**, 1315–1321 (1988).

[5] D. A. Kass, E. P. Shapiro, M. Kawaguchi, A. R. Capriotti, A. Scuteri, R. C. deGroof, and E. G. Lakatta, ''Improved arterial compliance by a novel advanced glycation end-product crosslink breaker,'' Circulation **104**, 1464–1470 (2001).

[6] A. J. Perejda, P. A. Abraham, W. H. Carnes, W. F. Coulson, and J. Uitto, ''Marfan's syndrome: structural, biochemical, and mechanical studies of the aortic media,'' J. Lab. Clin. Med. **106**, 376–383 (1985).

[7] Y. C. Fung, in *Biomechanics: Mechanical Properties of Living Tissues* (Springer-Verlag, New York, 1981), pp. 196–301.

[8] A. P. Avolio, S. G. Chen, R. P. Wang, C. L. Zhang, M. F. Li, and M. F. O'Rourke, ''Effects of aging on changing arterial compliance and left ventricular load in a northern Chinese urban community,'' Circulation **68**, 50–58 (1983).

[9] E. D. Lehmann, J. R. Parker, K. D. Hopkins, M. G. Taylor, and R. G. Gosling, ''Validation and reproducibility of pressure-corrected aortic distensibility measurements using pulse-wave-velocity Doppler ultrasound,'' J. Biomed. Eng. **15**, 221–228 (1993).

[10] V. V. Itskovich, K. A. Kraft, and D. Y. Fei, ''Rapid aortic wave velocity measurement with MR imaging,'' Radiology **219**, 551–557 (2001).

[11] C. Stefanadis, C. Stratos, H. Boudoulas, C. Kourouklis, and P. Toutouzas, ''Distensibility of the ascending aorta: comparison of invasive and noninvasive techniques in healthy men and in men with coronary artery disease,'' Eur. Heart J. **11**, 990–996 (1990).

[12] T. Länne, H. Stale, H. Bengtsson, D. Gustafsson, D. Bergqvist, B. Sonesson, H. Lecerof, and P. Dahl, ''Noninvasive measurement of diameter changes in the distal abdominal aorta in man,'' Ultrasound Med. Biol. **18**, 451–457 (1992).

[13] R. M. Lang, B. P. Cholley, C. Korcarz, R. H. Marcus, and S. G. Shroff, ''Measurement of regional elastic properties of the human aorta. A new application of transesophageal echocardiography with automated border detection and calibrated subclavian pulse tracings,'' Circulation **90**, 1875–1882 (1994).

[14] C. Stefanadis, C. Stratos, C. Vlachopoulos, S. Marakas, H. Boudoulas, I. Kallikazaros, E. Tsiamis, K. Toutouzas, L. Sioros, and P. Toutouzas, ''Pressure-diameter relation of the human aorta. A new method of determination by the application of a special ultrasonic dimension catheter,'' Circulation **92**, 2210–2219 (1995).

[15] S. Vulliemoz, N. Stergiopulos, and R. Meuli, ''Estimation of local aortic elastic properties with MRI,'' Magn. Reson. Med. **47**, 649–654 (2002).

[16] C. S. Hall, C. T. Nguyen, M. J. Scott, G. M. Lanza, and S. A. Wickline, ''Delineation of the extracellular determinants of ultrasonic scattering from elastic arteries,'' Ultrasound Med. Biol. **26**, 613–620 (2000).

[17] O. Ohtani, T. Ushiki, T. Taguchi, and A. Kikuta, ''Collagen fibrillar networks as skeletal frameworks: a demonstration by cell-maceration/scanning electron microscope method,'' Arch. Histol. Cytol. **51**, 249–261 (1988).

[18] T. Ushiki, ''Preserving the original architecture of elastin components in the formic acid-digested aorta by an alternative procedure for scanning electron microscopy,'' J. Electron. Microsc. (Tokyo) **41**, 60–63 (1992).

[19] V. A. Del Grosso and C. W. Mader, ''Speed of sound in pure water,'' J. Acoust. Soc. Am. **52**, 1442–1446 (1972).

[20] A. C. Simon, M. O'Rourke, and J. Levenson, ''Arterial distensibility and its effect on wave reflection and cardiac loading in cardiovascular disease,'' Coron. Artery Dis. **2**, 1111–1120 (1991).

[21] E. Picano, L. Landini, A. Distante, R. Sarnelli, A. Benassi, and A. L'Abbate, ''Different degrees of atherosclerosis detected by backscattered ultrasound: an in vitro study on fixed human aortic walls,'' J. Clin. Ultrasound **11**, 375–379 (1983).

[22] E. Picano, L. Landini, A. Distante, A. Benassi, R. Sarnelli, and A. L'Abbate, ''Fibrosis, lipids, and calcium in human atherosclerotic plaque. In vitro differentiation from normal aortic walls by ultrasonic attenuation,'' Circ. Res. **56**, 556–562 (1985).

[23] L. Landini, R. Sarnelli, E. Picano, and M. Salvadori, ''Evaluation of frequency dependence of backscatter coefficient in normal and atherosclerotic aortic walls,'' Ultrasound Med. Biol. **12**, 397–401 (1986).

[24] B. Barzilai, J. E. Saffitz, J. G. Miller, and B. E. Sobel, ''Quantitative ultrasonic characterization of the nature of atherosclerotic plaques in human aorta,'' Circ. Res. **60**, 459–463 (1987).

[25] S. L. Bridal, P. Fornes, P. Bruneval, and G. Berger, ''Parametric (integrated backscatter and attenuation) images constructed using backscattered radio frequency signals (25–56 MHz) from human aortae in vitro,'' Ultrasound Med. Biol. **23**, 215–229 (1997).

[26] R. K. Shepard, J. G. Miller, and S. A. Wickline, ''Quantification of atherosclerotic plaque composition in cholesterol-fed rabbits with 50-MHz acoustic microscopy,'' Arterioscler. Thromb. **12**, 1227–1234 (1992).

[27] Y. Saijo, H. Sasaki, H. Okawai, S. Nitta, and M. Tanaka, ''Acoustic properties of atherosclerosis of human aorta obtained with high-frequency ultrasound,'' Ultrasound Med. Biol. **24**, 1061–1064 (1998).

[28] D. Recchia, A. M. Sharkey, M. S. Bosner, N. T. Kouchoukos, and S. A. Wickline, ''Sensitive detection of abnormal aortic architecture in Marfan syndrome with high-frequency ultrasonic tissue characterization,'' Circulation **91**, 1036–1043 (1995).

[29] M. G. de Kroon, L. F. van der Wal, W. J. Gussenhoven, H. Rijsterborgh, and N. Bom, ''Backscatter directivity and integrated backscatter power of arterial tissue,'' Int. J. Card. Imaging **6**, 265–275 (1991).

[30] G. R. Lockwood, L. K. Ryan, J. W. Hunt, and F. S. Foster, ''Measurement of the ultrasonic properties of vascular tissues and blood from 35–65 MHz,'' Ultrasound Med. Biol. **17**, 653–666 (1991).

[31] C. T. Nguyen, C. S. Hall, and S. A. Wickline, ''Characterization of aortic microstructure with ultrasound: implications for mechanisms of aortic function and dissection,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **49**, 1561–1571 (2002).

[32] B. A. Auld, *Acoustic Fields and Waves in Solids*, 2nd ed. (Krieger, Malabar, FL, 1990).

[33] A. C. Burton, ''Relation of structure to function of the tissues of the wall of blood vessels,'' Physiol. Rev. **34**, 619–642 (1954).

[34] R. L. Armentano, E. I. Cabrera Fischer, J. G. Barra, J. A. Levenson, A. C. Simon, and R. H. Pichel, ''Single beat evaluation of circumferential aortic elastin elastic modulus in conscious dogs. Potential application in noninvasive measurements,'' Med. Prog. Technol. **20**, 91–99 (1994).

[35] Seymour Glagov, in *Blood Vessels and Lymphatics in Organ Systems*, edited by D. I. Abramson and P. B. Dobrin (Academic, Orlando, 1984), pp. 3–16.

[36] J. H. Rose, M. R. Kaufmann, S. A. Wickline, C. S. Hall, and J. G. Miller, ''A proposed microscopic elastic wave theory for ultrasonic backscatter from myocardial tissue,'' J. Acoust. Soc. Am. **97**, 656–668 (1995).

[37] B. K. Hoffmeister, S. M. Handley, E. D. Verdonk, S. A. Wickline, and J. G. Miller, ''Estimation of the elastic stiffness coefficient c13 of fixed tendon and fixed myocardium,'' J. Acoust. Soc. Am. **97**, 3171–3176 (1995).

[38] B. K. Hoffmeister, S. E. Gehr, and J. G. Miller, ''Anisotropy of the trans-

verse mode ultrasonic properties of fixed tendon and fixed myocardium,'' J. Acoust. Soc. Am. **99**, 3826–3836 (1996).
[39] B. K. Hoffmeister, S. M. Handley, S. A. Wickline, and J. G. Miller, ''Ultrasonic determination of the anisotropy of Young's modulus of fixed tendon and fixed myocardium,'' J. Acoust. Soc. Am. **100**, 3933–3940 (1996).
[40] J. V. Geleskie and K. K. Shung, ''Further studies on acoustic impedance of major bovine blood vessel walls,'' J. Acoust. Soc. Am. **71**, 467–470 (1982).
[41] J. C. Park, R. J. Siegel, and L. L. Demer, ''Effect of calcification and formalin fixation on in vitro distensibility of human femoral arteries,'' Am. Heart J. **125**, 344–349 (1993).
[42] J. C. Bamber, C. R. Hill, J. A. King, and F. Dunn, ''Ultrasonic propagation through fixed and unfixed tissues,'' Ultrasound Med. Biol. **5**, 159–165 (1979).
[43] A. F. van der Steen, M. H. Cuypers, J. M. Thijssen, and P. C. de Wilde, ''Influence of histochemical preparation on acoustic parameters of liver tissue: a 5-MHz study,'' Ultrasound Med. Biol. **17**, 879–891 (1991).
[44] C. S. Hall, C. L. Dent, M. J. Scott, and S. A. Wickline, ''High-frequency ultrasound detection of the temporal evolution of protein cross linking in myocardial tissue,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **47**, 1051–1058 (2000).
[45] E. G. Tickner and A. H. Sacks, ''A theory for the static elastic behavior of blood vessels,'' Biorheology **4**, 151–168 (1967).
[46] D. J. Patel, J. S. Janicki, and T. E. Carew, ''Static anisotropic elastic properties of the aorta in living dogs,'' Circ. Res. **25**, 765–779 (1969).
[47] H. W. Weizsacker and T. D. Kampp, ''Passive elastic properties of the rat aorta,'' Biomed. Tech. **35**, 224–234 (1990).
[48] J. Zhou and Y. C. Fung, ''The degree of nonlinearity and anisotropy of blood vessel elasticity,'' Proc. Natl. Acad. Sci. U.S.A. **94**, 14255–14260 (1997).

Lamb wave propagation in elastic plates coated with viscoelastic materials

F. Simonetti
Department of Mechanical Engineering, Imperial College, London SW7 2AZ, United Kingdom

(Received 13 October 2003; revised 14 February 2004; accepted 15 February 2004)

This paper addresses the effects of attenuative coatings on the dispersion characteristics of Lamb wave propagation in elastic plates. The topology of phase velocity and guided wave attenuation spectra is analyzed as a function of the coating internal damping (longitudinal and shear bulk attenuations) and it is proved that in contrast to elastic plates, all modes are propagating albeit with large attenuation in some cases. An energy-based correspondence between the dispersion of the attenuative bilayer and that of a related elastic bilayer is derived in order to investigate separately the effects of the longitudinal and shear bulk attenuations on the attenuation of the guided modes. It is shown that at low frequency the guided wave attenuation is only slightly affected by the longitudinal bulk attenuation, while the contribution of the shear bulk attenuation is substantial. The attenuation characteristics of shear horizontal modes are compared with those of Lamb modes in order to identify the mode and the frequencies which result in minimum guided wave attenuation.
 [DOI: 10.1121/1.1695011]

PACS numbers: 43.35.Mr, 43.35.Zc [YHB] Pages: 2041–2053

I. INTRODUCTION

The nondestructive testing of large metallic structures by using ultrasonic guided waves is now widely employed. The technique is particularly advantageous since the acoustic signal is guided by the structure so a large length of one-dimensional structures or area of two-dimensional structures can be inspected. In both cases, the detection of flaws is achieved by exciting the acoustic signal at one location of the structure and by looking at the echoes produced by the defects. Moreover, in contrast with traditional techniques which are based on the propagation of longitudinal and shear bulk waves, a broad range of guided waves can be excited. Depending on the geometry of the structure, on the position of the defects, and on the surrounding environment, a mode can be selected in order to obtain the highest sensitivity. The use of either extensional[1–13] or distortional[14,15] modes for inspecting small diameter heat exchangers, large diameter pipelines, and storage tank plates has been studied extensively and screening devices have been designed and commercialized for routine testing.[16,17]

However, guided waves only provide a satisfactory alternative to traditional methods if the acoustic signal can propagate for a sufficiently large distance while maintaining an energy level higher than the background noise. In one-dimensional lossless structures, the signal propagates without being attenuated, while in two-dimensional structures, the signal decays approximately with the square root of the distance.[18] However, while internal losses in metals are negligible at the low frequencies used, the presence of attenuative media in combination with the waveguide can dramatically attenuate the signal. Since a significant proportion of industrial structures are coated or embedded in other media, the assessment of the attenuation characteristics of both extensional and distortional modes becomes a major issue.

A particular concern arises from the presence of attenuative coatings, such as in the case of bitumen coated pipelines, and the purpose of this paper is to address the attenuation characteristics of extensional guided waves propagating in metallic plates coated with viscoelastic layers. This geometry is also representative of hollow cylinders with large diameter to wall thickness ratio.

There have been few studies on this subject. Jones[19] and Laperre[20] considered the propagation of Lamb waves in bilayered elastic plates but they did not explore the effects of internal losses. On the other hand, the influence of a Newtonian viscous fluid layer on the dispersion of Lamb waves in metallic plates was investigated by Zhu[21] and by Nayfeh.[22] However, this approach cannot be used for viscoelastic coatings as the shear elasticity, which is not considered in the fluid model, plays a crucial role in the attenuation characteristics. More recently, Drinkwater *et al.*[23] studied the propagation of the A_0 and S_0 Lamb modes in a glass plate compressively loaded with a highly attenuative elastomer.

In this paper, the internal losses in the coating are modeled according to the theory of linear viscoelasticity, which is also the model implemented in the software DISPERSE[24] used for validating some of the results of this work. The features of the dispersion of Lamb modes are studied for different regimes of material damping in order to identify the conditions which result in minimum guided wave attenuation and therefore maximize the range of inspection. Moreover, so as to provide a comprehensive investigation of all the bilayer modes the attenuation characteristics of shear horizontal (SH) modes[25] are revisited and compared to those of Lamb waves.

II. SECULAR EQUATIONS FOR A BILAYER

Here, an elastic, isotropic and homogeneous metallic plate of infinite extent, coated with a viscoelastic layer, is considered (Fig. 1). The layer is modeled according to the

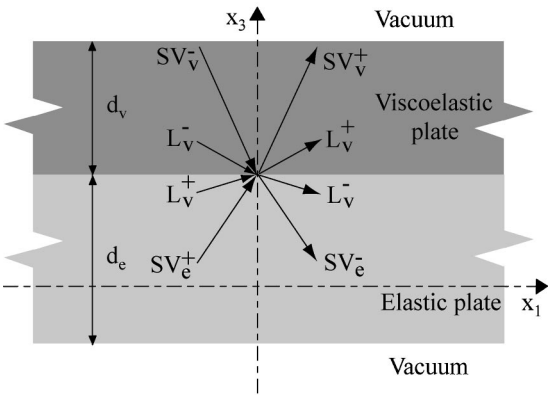


FIG. 1. Partial waves diagram.

theory of linear viscoelasticity for isotropic and homogeneous media which, in the frequency domain, leads to linear equations of motion[26]

$$(\tilde{\lambda}+\tilde{\mu})\nabla(\nabla\cdot\mathbf{u})+\tilde{\mu}\nabla^2\mathbf{u}+\rho\omega^2\mathbf{u}=\mathbf{0}, \tag{1}$$

where $\mathbf{u}$ is the Fourier transformed displacement vector, ρ is the density, ∇ is the three-dimensional differential operator and ω is the angular frequency. The complex frequency-dependent functions $\tilde{\lambda}$ and $\tilde{\mu}$ are related to the relaxation functions of the material $\lambda(t)$ and $\mu(t)$,

$$\tilde{\lambda}=\lambda_0+i\omega\int_0^\infty \lambda(t)e^{i\omega t}\,dt, \tag{2}$$

$$\tilde{\mu}=\mu_0+i\omega\int_0^\infty \mu(t)e^{i\omega t}\,dt, \tag{3}$$

where λ_0 and μ_0 are the asymptotic values of the relaxation curves.

It can be shown that for an unbounded space the solution to (1) can be thought of as a superposition of longitudinal, L, and shear, S, harmonic waves of the form

$$\mathbf{u}_{\mathrm{L,S}}=A_{\mathrm{L,S}}e^{i\mathbf{k}_{\mathrm{L,S}}\cdot\mathbf{x}}\mathbf{n}_{\mathrm{L,S}}, \tag{4}$$

where $A_{\mathrm{L,S}}$ are arbitrary complex constants, $\mathbf{x}$ is the position vector, $\mathbf{n}_{\mathrm{L,S}}$ are the polarization directions, and $\mathbf{k}_{\mathrm{L,S}}$ are the complex wavenumber vectors, which satisfy the secular equation

$$\mathbf{k}_{\mathrm{L,S}}\cdot\mathbf{k}_{\mathrm{L,S}}=\frac{\omega^2}{a_{\mathrm{L,S}}^2}, \tag{5}$$

in which the complex sound velocities, $a_{\mathrm{L,S}}$, are defined according to

$$a_{\mathrm{L}}^2=\frac{\tilde{\lambda}+2\tilde{\mu}}{\rho}, \tag{6}$$

$$a_{\mathrm{S}}^2=\frac{\tilde{\mu}}{\rho}, \tag{7}$$

which can also be expressed as

$$a_{\mathrm{L,S}}=\frac{c_{\mathrm{L,S}}}{1+i\alpha_{\mathrm{L,S}}/2\pi}, \tag{8}$$

where $c_{\mathrm{L,S}}$ are the material bulk velocities and $\alpha_{\mathrm{L,S}}$ are the bulk attenuations in Nepers per wavelength.

Let us now consider the stress associated with a bulk wave, either shear or longitudinal, whose displacement field is described by (4). It is known that the Lamé constitutive equations

$$\sigma_{ij}=\tilde{\lambda}\delta_{ij}\varepsilon_0+2\tilde{\mu}\varepsilon_{ij}, \tag{9}$$

and the compatibility equations,

$$\varepsilon_{ij}=\frac{1}{2}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right), \tag{10}$$

$$\varepsilon_0=\varepsilon_{11}+\varepsilon_{22}+\varepsilon_{33}, \tag{11}$$

allow the stress to be derived from the displacement field. Moreover, since the Lamé moduli can be expressed as a function of the complex velocities, $a_{\mathrm{L,S}}$, and the density by inverting (6) and (7), it follows that the stress tensor can be regarded as the product of the density, a suitable tensor T_{ij} which depends on the bulk velocities only and the amplitude, A, of the bulk wave

$$\sigma_{ij}=\rho T_{ij}A, \tag{12}$$

$$T_{ij}A=(a_{\mathrm{L}}^2-2a_{\mathrm{S}}^2)\delta_{ij}\varepsilon_0+2a_{\mathrm{S}}^2\varepsilon_{ij}. \tag{13}$$

The displacement field of an SH wave propagating in a bilayered plate results from the superposition of the fields associated with an upward and a downward shear bulk wave per layer,[27] each bulk wave propagating in the $x_2=0$ plane and being polarized parallel to $\hat{x}_2$. The secular equation can be found by imposing continuity of stress and displacement at the interface between the two layers, and the zero traction condition at the free boundaries of the bilayer. This leads to a system of four equations in the four unknown bulk shear wave amplitudes, $A^{\mathrm{SH}_{v,e}^{\pm}}$

$$\begin{bmatrix} \mathrm{T}_{3,2_{top}}^{\mathrm{SH}_v^+} & \mathrm{T}_{3,2_{top}}^{\mathrm{SH}_v^-} & 0 & 0 \\ u_{2_{int}}^{\mathrm{SH}_v^+} & u_{2_{int}}^{\mathrm{SH}_v^-} & -u_{2_{int}}^{\mathrm{SH}_e^+} & -u_{2_{int}}^{\mathrm{SH}_e^-} \\ \frac{\rho_v}{\rho_e}\mathrm{T}_{3,2_{int}}^{\mathrm{SH}_v^+} & \frac{\rho_v}{\rho_e}\mathrm{T}_{3,2_{int}}^{\mathrm{SH}_v^-} & -\mathrm{T}_{3,2_{int}}^{\mathrm{SH}_e^+} & -\mathrm{T}_{3,2_{int}}^{\mathrm{SH}_e^-} \\ 0 & 0 & \mathrm{T}_{3,2_{bot}}^{\mathrm{SH}_e^+} & \mathrm{T}_{3,2_{bot}}^{\mathrm{SH}_e^-} \end{bmatrix} \begin{bmatrix} A^{\mathrm{SH}_v^+} \\ A^{\mathrm{SH}_v^-} \\ A^{\mathrm{SH}_e^+} \\ A^{\mathrm{SH}_e^-} \end{bmatrix} = 0, \tag{14}$$

where the subscripts v and e refer to the viscoelastic and elastic layers, respectively, while the signs $\pm$ indicate the upward and downward directions of the bulk waves, respectively. Since the displacement is polarized parallel to $\hat{x}_2$ and is constant along this direction, it follows that the stress on the planes $x_3=\mathrm{const}$ is parallel to $\hat{x}_2$. The first row in the equation (14) gives the traction-free condition on the top of the bilayer, the second and third provide the continuity of displacement and stress at the interface, respectively, and the last gives the traction-free condition at the bottom of the bilayer. The system (14) admits nontrivial solutions only if the determinant of the matrix vanishes. This condition provides the secular equation for SH waves in bilayered plates.

The phase velocity of the resulting guided wave is

$$C_{ph}=\frac{\omega}{\mathrm{Re}\{\mathbf{k}_{\|}\}}, \tag{15}$$

where $\mathbf{k}_{\|}$ is the projection of the wavenumber along the interface ($\mathbf{k}_{\|}$ is an invariant for all the partial waves as Snell's law holds). Moreover, the imaginary part of $\mathbf{k}_{\|}$ gives the guided wave attenuation, ζ. Furthermore, the Fourier transformed displacement field has the form

$$u_2(x_1x_3,\omega)=h_2(x_3)e^{-\zeta x_1}e^{-i(\omega/C_{ph})x_1}, \tag{16}$$

where the function h_2 is the mode shape.

Let us now consider the propagation of Lamb waves, which originate from the superposition of one pair of longitudinal bulk waves, $L^{\pm}$, and one pair of shear bulk waves, $SV^{\pm}$, per layer[28] (Fig. 1), the bulk waves being polarized parallel to the plane $x_2=0$. The boundary conditions, which are the same as in the case of SH waves, lead to the following secular equation:

$$\det\begin{bmatrix}
T^{SV_v^+}_{3,1_{top}} & T^{SV_v^-}_{3,1_{top}} & T^{L_v^+}_{3,1_{top}} & T^{L_v^-}_{3,1_{top}} & 0 & 0 & 0 & 0\\
T^{SV_v^+}_{3,3_{top}} & T^{SV_v^-}_{3,3_{top}} & T^{L_v^+}_{3,3_{top}} & T^{L_v^-}_{3,3_{top}} & 0 & 0 & 0 & 0\\
u^{SV_v^+}_{1_{int}} & u^{SV_v^-}_{1_{int}} & u^{L_v^+}_{1_{int}} & u^{L_v^-}_{1_{int}} & -u^{SV_e^+}_{1_{int}} & -u^{SV_e^-}_{1_{int}} & -u^{L_e^+}_{1_{int}} & -u^{L_e^-}_{1_{int}}\\
u^{SV_v^+}_{3_{int}} & u^{SV_v^-}_{3_{int}} & u^{L_v^+}_{3_{int}} & u^{L_v^-}_{3_{int}} & -u^{SV_e^+}_{3_{int}} & -u^{SV_e^-}_{3_{int}} & -u^{L_e^+}_{3_{int}} & -u^{L_e^-}_{3_{int}}\\
\frac{\rho_v}{\rho_e}T^{SV_v^+}_{3,1_{int}} & \frac{\rho_v}{\rho_e}T^{SV_v^-}_{3,1_{int}} & \frac{\rho_v}{\rho_e}T^{L_v^+}_{3,1_{int}} & \frac{\rho_v}{\rho_e}T^{L_v^-}_{3,1_{int}} & -T^{SV_e^+}_{3,1_{int}} & -T^{SV_e^-}_{3,1_{int}} & -T^{L_e^+}_{3,1_{int}} & -T^{L_e^-}_{3,1_{int}}\\
\frac{\rho_v}{\rho_e}T^{SV_v^+}_{3,3_{int}} & \frac{\rho_v}{\rho_e}T^{SV_v^-}_{3,3_{int}} & \frac{\rho_v}{\rho_e}T^{L_v^+}_{3,3_{int}} & \frac{\rho_v}{\rho_e}T^{L_v^-}_{3,3_{int}} & -T^{SV_e^+}_{3,3_{int}} & -T^{SV_e^-}_{3,3_{int}} & -T^{L_e^+}_{3,3_{int}} & -T^{L_e^-}_{3,3_{int}}\\
0 & 0 & 0 & 0 & T^{SV_e^+}_{3,1_{bot}} & T^{SV_e^-}_{3,1_{bot}} & T^{L_e^+}_{3,1_{bot}} & T^{L_e^-}_{3,1_{bot}}\\
0 & 0 & 0 & 0 & T^{SV_e^+}_{3,3_{bot}} & T^{SV_e^-}_{3,3_{bot}} & T^{L_e^+}_{3,3_{bot}} & T^{L_e^-}_{3,3_{bot}}
\end{bmatrix}=0, \tag{17}$$

where the first two rows refer to the traction-free condition on the top of the bilayer, the third and fourth rows account for the continuity of the displacement at the bilayer interface, the fifth and sixth come from the stress continuity at the interface, and the last two provide the traction-free condition at the bottom of the bilayer. In this case, the resulting displacement field can be expressed as

$$u_1(x_1,x_3,\omega)=h_1(x_3)e^{-\zeta x_1}e^{-i(\omega/C_{ph})x_1}, \tag{18}$$

$$u_3(x_1,x_3,\omega)=h_3(x_3)e^{-\zeta x_1}e^{-i(\omega/C_{ph})x_1}. \tag{19}$$

III. GUIDED WAVE ATTENUATION

The purpose of this section is to provide a relationship between the guided wave attenuation, the strain energy in the viscoelastic layer, and the acoustic properties of the bilayer. For the case of SH waves, it has already been shown[25] that by considering a volume, V, of unit width in the $\hat{x}_2$ direction, and with height equal to the thickness of the bilayer (see Fig. 2), the guided wave attenuation, ζ, can be related to the average dissipated power, P_d, in the volume and the average in-plane power flow per unit width, $\langle P\rangle$,

$$\zeta=\frac{dP_d/dx_1}{2\langle P\rangle}. \tag{20}$$

The same relationship holds in the case of Lamb waves. This can be proved by means of the complex acoustic Poynting's vector theorem[29] which for a volume, V, bounded by the closed surface, S, can be written as

$$\oint_S \mathbf{P}\cdot\mathbf{n}\,dS-i\omega(U-K)+P_d=0, \tag{21}$$

where U and K are the peak strain and kinetic energies stored in V, respectively. P_d is the average power dissipated over one cycle within V, $\mathbf{n}$ is the outward normal to S, and $\mathbf{P}$ is the Poynting vector whose jth component, in tensor notation, is defined as

$$P_j=-i\omega\frac{u_i^*\sigma_{ij}}{2}, \tag{22}$$

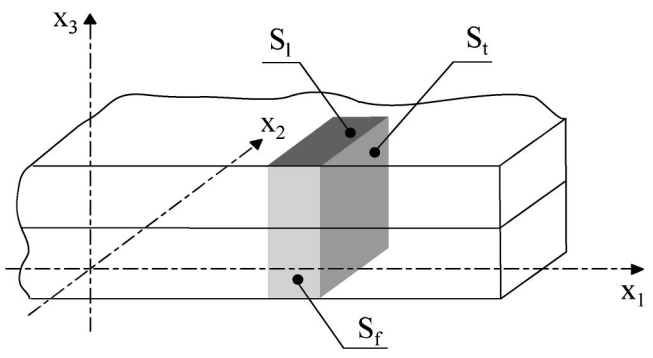


FIG. 2. Volume considered for the energy balance.

in which the superscript $*$ refers to the complex conjugate of the quantity and σ is the local stress tensor. The real part of the complex Poynting vector flux is the average net power flow into S over one cycle.[29] Moreover, by considering the real part of Eq. (21) it follows that the dissipated power, P_d, equals in magnitude the average power flow through S (since U, K, and P_d are real). The net flux into the rectangular parallelepiped shown in Fig. 2 is given by the summation of the flows through the faces of the parallelepiped. However, the flux through each lateral surface perpendicular to $\hat{x}_3$, S_l, is zero due to the traction-free condition. Also the flow across the faces perpendicular to $\hat{x}_2$, S_f, vanishes as the only nonzero stress is σ_{22} which is perpendicular to the displacement field. By substituting (18) and (19) into (22) the flux through a transversal surface perpendicular to $\hat{x}_1$, S_t, at the position x_1 can be written as

$$P_{x_1}=\frac{1}{2}\omega e^{-2\zeta x_1}\int_{d_e+d_v}(\mathbf{k}_{\|}((\tilde{\lambda}+2\tilde{\mu})|h_1|^2+\tilde{\mu}|h_3|^2) + i(\tilde{\lambda}h_3'h_1^*+\tilde{\mu}h_1'h_3^*)dx_3. \tag{23}$$

As a consequence, the total net flow into the parallelepiped is the difference between the fluxes calculated at the two transversal surfaces. Moreover, by considering the real part of (21), the power balance can be written as

$$\Delta\ \mathrm{Re}\{P_{x_1}\}+P_d=0. \tag{24}$$

By substituting (23) into (24) and differentiating with respect to x_1, and by considering that the real part of P_{x_1} is the average in-plane power flow, $\langle P\rangle$, one obtains (20). Moreover, Eq. (24) has a fundamental implication for the nonpropagating modes. By definition,[30] a mode is said to be nonpropagating if $\langle P\rangle=0$ (the mode does not carry energy). As a consequence, if there were nonpropagating modes, from (24), the system could vibrate without dissipating energy which is physically impossible. It follows that all the modes have to be propagating as has already been shown for SH waves.[25]

Generally speaking, at a prescribed point of a viscoelastic medium, the average dissipated power per unit volume can be related to the peak strain energy per unit volume.[29] In particular, for SH waves the peak strain energy per unit volume, e_s, is

$$e_s=2\tilde{\mu}_r^v(\varepsilon_{12}\varepsilon_{12}^*+\varepsilon_{23}\varepsilon_{23}^*), \tag{25}$$

in which the subscript r refers to the real part of the quantity. On the other hand, the average dissipated power per unit volume, p_d, is given by

$$p_d=2\omega\tilde{\mu}_i^v(\varepsilon_{12}\varepsilon_{12}^*+\varepsilon_{23}\varepsilon_{23}^*), \tag{26}$$

where the subscript i refers to the imaginary part of the quantity. Since for the bilayer all the field functions are constant along the $\hat{x}_2$ direction it follows that the total power dissipated in V has to satisfy

$$\frac{dP_d}{dx_1}=\int_{d_v}p_d\,dx_3=\omega\frac{\mu_i^v}{\mu_r^v}\int_{d_v}e_d\,dx_3=\omega\frac{\mu_i^v}{\mu_r^v}\frac{dE}{dx_1}, \tag{27}$$

where E is the peak strain energy of the portion of the viscoelastic layer contained in V. As a result, the guided wave attenuation can be expressed as

$$\zeta=\frac{1}{2}\omega\frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v}Q_E, \tag{28}$$

where

$$Q_E=\frac{dE/dx_1}{\langle P\rangle} \tag{29}$$

is the first energy factor already introduced by the authors.[25] Note that the energy factor does not correspond to the slowness as it involves the strain energy contained in the viscoelastic layer only.

Similarly, in the case of Lamb waves the peak strain energy and the dissipated power per unit volume, after some manipulations, can be expressed as

$$e_s=\tfrac{1}{2}(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v)\varepsilon_0\varepsilon_0^*+2\tilde{\mu}_r^v(\varepsilon_{13}\varepsilon_{13}^*-\varepsilon_{11}\varepsilon_{33}^*), \tag{30}$$

$$p_d=\tfrac{1}{2}\omega(\tilde{\lambda}_i^v+2\tilde{\mu}_i^v)\varepsilon_0\varepsilon_0^*+2\omega\tilde{\mu}_i^v(\varepsilon_{13}\varepsilon_{13}^*-\varepsilon_{11}\varepsilon_{33}^*), \tag{31}$$

which leads to

$$p_d=\omega\left(\frac{\tilde{\lambda}_i^v+2\tilde{\mu}_i^v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\delta+\frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v}\gamma\right), \tag{32}$$

where

$$\delta=\tfrac{1}{2}(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v)\varepsilon_0\varepsilon_0^*, \tag{33}$$

$$\gamma=2\tilde{\mu}_r^v(\varepsilon_{13}\varepsilon_{13}^*-\varepsilon_{11}\varepsilon_{33}^*). \tag{34}$$

It is interesting to note that δ only depends on the displacement field resulting from the superposition of the two longitudinal bulk waves propagating in the viscoelastic layer. In order to show this, it is sufficient to consider that the strain tensor due to the superposition of two pairs of longitudinal and shear bulk waves is equivalent to the sum of the strain tensors of each bulk wave taken separately. Since for a bulk shear wave the first strain invariant vanishes, it follows that ε_0 is the sum of the invariants of the two longitudinal bulk waves only. However, this does not necessarily mean that δ coincides with the total energy carried by the longitudinal waves. In contrast with δ, γ depends on the presence of both shear and longitudinal partial waves in the viscoelastic layer. As a consequence, the δ energy of a mode whose shear partial waves are dominated by the longitudinal ones is not necessarily larger than γ as will be shown later in Sec. V.

As in the case of SH waves, the derivative of the dissipated power can be expressed as

$$\frac{dP_d}{dx_1}=\int_{d_v}p_d\,dx_3=\omega\left(\frac{\tilde{\lambda}_i^v+2\tilde{\mu}_i^v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\int_{d_v}\delta\,dx_3+\frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v}\int_{d_v}\gamma\,dx_3\right) =\omega\left(\frac{\tilde{\lambda}_i^v+2\tilde{\mu}_i^v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\frac{d\Delta}{dx_1}+\frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v}\frac{d\Gamma}{dx_1}\right), \tag{35}$$

where Δ and Γ are the integrals of δ and γ over the portion of the viscoelastic layer contained in V. By substituting (35) into (20) the guided wave attenuation is given by

$$\zeta = \frac{1}{2}\omega\left(\frac{\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v}{\tilde{\lambda}_r^v + 2\tilde{\mu}_r^v} Q_\Delta + \frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v} Q_\Gamma\right), \tag{36}$$

where the two energy factors Q_Δ and Q_Γ are defined as

$$Q_\Delta = \frac{d\Delta/dx_1}{\langle P\rangle}, \tag{37}$$

$$Q_\Gamma = \frac{d\Gamma/dx_1}{\langle P\rangle}. \tag{38}$$

Equations (28) and (36) show that the guided wave attenuation can always be expressed as a linear combination of the energy factors, each energy factor being related to the strain energy contained in the viscoelastic layer when unit power flows through the cross section of the bilayer. Moreover, Eqs. (28) and (36) allow the guided wave attenuation dispersion to be derived from the dispersion of a suitable equivalent elastic bilayer. In order to show this, the Maclaurin expansion of the function $\zeta(\tilde{\mu}_i^v)$ in the case of SH waves[25] or $\zeta(\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v, \tilde{\mu}_i^v)$ for Lamb waves is considered. In particular, by taking the expansions up to the first order ζ results in

$$\zeta_{\mathrm{SH}} \simeq \frac{1}{2}\omega\frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v} Q_E\big|_{\tilde{\mu}_i^v = 0}, \tag{39}$$

$$\zeta_{\mathrm{Lamb}} \simeq \frac{1}{2}\omega\left(\frac{\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v}{\tilde{\lambda}_r^v + 2\tilde{\mu}_r^v} Q_\Delta\bigg|_{\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v = 0} + \frac{\tilde{\mu}_i^v}{\tilde{\mu}_r^v} Q_\Gamma\bigg|_{\tilde{\mu}_i^v = 0}\right), \tag{40}$$

which, by considering that for low material damping the bulk attenuations in nepers per meter take the form[29]

$$\alpha_{\mathrm{S}} \simeq \pi\frac{\tilde{\mu}_i}{\tilde{\mu}_r}, \tag{41}$$

$$\alpha_{\mathrm{L}} \simeq \pi\frac{\tilde{\lambda}_i + 2\tilde{\mu}_i}{\tilde{\lambda}_r + 2\tilde{\mu}_r}, \tag{42}$$

allow the guided wave attenuation to be written as

$$\zeta_{\mathrm{SH}} \simeq f\alpha_{\mathrm{S}} Q_E\big|_{\tilde{\mu}_i^v = 0}, \tag{43}$$

$$\zeta_{\mathrm{Lamb}} \simeq f\alpha_{\mathrm{L}} Q_\Delta\big|_{\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v = 0} + f\alpha_{\mathrm{S}} Q_\Gamma\big|_{\tilde{\mu}_i^v = 0}. \tag{44}$$

These equations clearly illustrate how splitting the energy into two parts, δ and γ, allows the guided wave attenuation to be thought of as a superposition of the effects of the two material bulk attenuation coefficients (longitudinal and shear) taken separately. In particular, the δ energy which depends on the energy carried by the longitudinal partial waves only (it would be zero if there were not longitudinal partial waves) provides, through the energy factor Q_Δ, the part of guided wave attenuation due the longitudinal bulk attenuation coefficient. On the other hand, the γ energy gives, through the energy factor Q_Γ, the part of guided wave attenuation due to the bulk shear attenuation coefficient. It has to be emphasized that (44) does not imply that the guided wave attenuation can be expressed as a linear superposition of the attenuation due to longitudinal and shear partial waves considered separately as this would violate the nonlinearity of the attenuation phenomenon (attenuation depends on energy loss, therefore it cannot be linear). The effect of such a nonlinearity is contained in the definition of Q_Γ which depends on the energy carried by the longitudinal and shear partial waves.

In Eqs. (43) and (44), the energy factors are calculated by considering the modes of an artificial bilayer in which the viscoelastic layer has been replaced with an ideal layer whose Lamé moduli are purely real and equal to the real parts of the viscoelastic layer moduli [i.e., $\mu_{\mathrm{art}} = \mathrm{Re}\{\tilde{\mu}(\omega)\}$ and $\lambda_{\mathrm{art}} = \mathrm{Re}\{\tilde{\lambda}(\omega)\}$].

At a given frequency, the artificial bilayer behaves as if it was perfectly elastic. However, its response in the frequency domain differs from that of an elastic bilayer since the ideal layer is not strictly elastic due to the frequency dependence of its Lamé moduli [$\mu_{\mathrm{art}}(\omega)$ and $\lambda_{\mathrm{art}}(\omega)$]. Nevertheless, it can be observed that the artificial moduli μ_{art} and λ_{art} can always be expressed as sum of a constant term (μ_{equ} or λ_{equ}) and a frequency-dependent function [$\mu_{\mathrm{var}}(\omega)$ or $\lambda_{\mathrm{var}}(\omega)$]:

$$\mu_{\mathrm{art}} = \mu_{\mathrm{equ}} + \mu_{\mathrm{var}}(\omega), \tag{45}$$

$$\lambda_{\mathrm{art}} = \lambda_{\mathrm{equ}} + \lambda_{\mathrm{var}}(\omega). \tag{46}$$

As a consequence, the dispersion of the artificial bilayer can be thought of as a superposition of the dispersion of an ''equivalent'' elastic bilayer obtained by setting the Lamé moduli to be real and constant with frequency ($\mu_{\mathrm{art}} = \mu_{\mathrm{equ}}$ and $\lambda_{\mathrm{art}} = \lambda_{\mathrm{equ}}$) and a perturbation term which accounts for the frequency dependence of the artificial moduli (i.e., μ_{var} and λ_{var}).[25] However, the first-order approximation holds under the hypothesis of low material damping which implies that the dispersion of the real part of the Lamé moduli is negligible. As a consequence, the perturbation factor can also be neglected.

It has to be emphasized that the energy factors associated with a mode of the equivalent elastic bilayer are singular in the nonpropagating region of the mode since the in-plane power flow vanishes. As a consequence, the Maclaurin expansions cannot be employed. On the other hand, it is possible to relate the guided wave attenuation to the time derivative of the mechanical energy of the equivalent elastic bilayer by considering the imaginary part of (21):

$$\Delta(P_{x_1})_i - \omega(U - K) = 0. \tag{47}$$

By differentiating this expression with respect to x_1 and by taking into account (23), a new expression for the guided wave attenuation is obtained,

$$\zeta = \omega\frac{(d/dx_1)(U - K)}{2(P_{x_1})_i}, \tag{48}$$

which relates the mechanical energy decay to the guided wave attenuation. As a consequence, the guided wave attenuation can be expressed as a function of a second energy factor, Π,

$$\zeta = \tfrac{1}{2}\omega\Pi, \tag{49}$$

where Π is defined according to

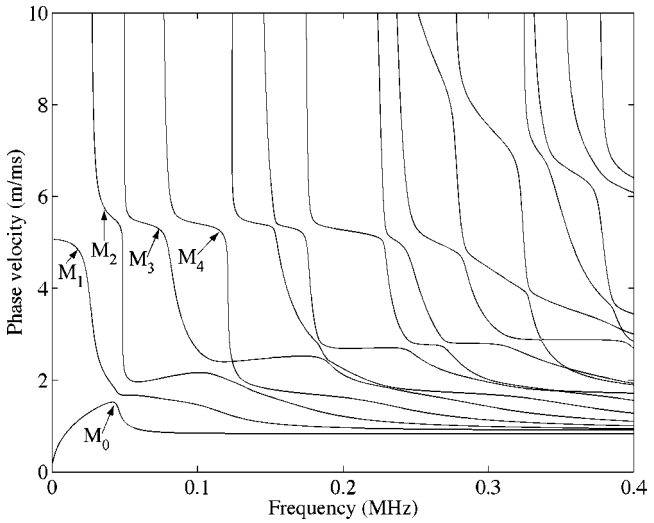


FIG. 3. Lamb wave phase velocity dispersion curves for the elastic bilayer described in Table I.

$$\Pi = \frac{(d/dx_1)(U-E)}{(P_{x_1})_i}. \tag{50}$$

The guided wave attenuation, in the case of low loss materials, can be approximated by considering the Maclaurin series of the function $\zeta(\tilde{\mu}_i^v)$ in the case of SH waves or $\zeta(\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v, \tilde{\mu}_i^v)$ for Lamb waves up to zeroth order

$$\zeta \simeq \tfrac{1}{2}\omega\overline{\Pi}, \tag{51}$$

where $\overline{\Pi}$ is calculated at $\tilde{\mu}_i^v = 0$ in the case of SH waves or $\tilde{\lambda}_i^v + 2\tilde{\mu}_i^v = \tilde{\mu}_i^v = 0$ for Lamb waves. Since ζ is the projection of the imaginary part of the wavenumber along the interface, $\overline{\Pi}$ can be obtained from the value of the imaginary part of the wavenumber of the nonpropagating mode through (51):

$$\overline{\Pi} = \frac{2}{\omega}(\mathbf{k}_{\|})_i. \tag{52}$$

It follows that, under the zeroth-order approximation, the guided wave attenuation in the attenuative system corresponds to the projection along the bilayer interface of the wavenumber of the nonpropagating mode of the corresponding elastic system.

IV. ELASTIC BILAYER

As has been shown in the previous section, the guided wave attenuation dispersion can be derived from the energy factors of the equivalent elastic bilayer. For this reason this section is devoted to the study of the dispersion of elastic bilayers. Figure 3 shows the Lamb wave dispersion curves obtained by solving Eq. (17) when the material damping is neglected, the bilayer acoustic properties and the geometry being summarized in Table I. It can be seen that these curves are considerably different from those of a metallic plate loaded with an inviscid fluid layer. In the latter case, due to the absence of the shear stiffness in the liquid layer, the dispersion curves are only slightly different from those of the free metallic plate.[31]

TABLE I. Material parameters used for the analysis.

	c_S (m/ms)	c_L (m/ms)	ρ (kg/m^3)	Thickness (mm)
Elastic plate	0.90	1.70	1250	9.00
Metallic plate	3.26	5.96	7930	8.00

The SH and Lamb waves of the bilayer originate from the interaction between the modes of the free metallic plate and those of the equivalent viscoelastic layer if it were rigidly clamped at the bilayer interface. In order to show this, it can be observed that for both SH and Lamb waves the matrices in (14) and (17) can be partitioned into four square matrices and the characteristic equation can be generalized according to

$$\det\begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} = 0, \tag{53}$$

where the matrices $\mathbf{C}_{ij}$ have dimension 2×2 for SH waves and 4×4 for Lamb waves. Let us consider the limit as the stiffness and the density of the metallic plate go to infinity while keeping the complex bulk shear and longitudinal velocities constant. Since the displacements and the quantities T_{ij} depend only on the two bulk velocities, the matrices $\mathbf{C}_{ij}$ remain unchanged except for $\mathbf{C}_{21}$ which vanishes (since $\rho_v/\rho_e \to 0$). As a consequence, for this asymptotic condition, the characteristic equation becomes[32]

$$\det\begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{0} & \mathbf{C}_{22} \end{bmatrix} = \det(\mathbf{C}_{11})\det(\mathbf{C}_{22}) = 0, \tag{54}$$

which is satisfied if either $\det(\mathbf{C}_{11}) = 0$ or $\det(\mathbf{C}_{22}) = 0$. In particular, the first condition corresponds to wave propagation in the equivalent viscoelastic layer if it were rigidly clamped at the interface. This is due to the fact that $\mathbf{C}_{11}$ is a function of the stress at the top of the bilayer and the displacement of the equivalent viscoelastic layer at the interface. The second condition coincides with the characteristic equation for the free elastic plates as $\mathbf{C}_{22}$ is a function of the stress on the two faces of the elastic plate. As a consequence, the asymptotic solution can be divided into two families of modes. The first family is represented by the free metallic plate modes while the second one corresponds to the modes of the clamped-free equivalent viscoelastic layer. As an example, Fig. 4 shows the two asymptotic families in the case of Lamb waves. For a finite value of the density and stiffness of the metallic plate, the bilayer mode trajectories jump from one asymptotic mode to another, as shown in Fig. 5. For instance, let us consider the path of the bilayer mode M_3. Its cutoff frequency occurs close to that of the M_2' mode of the equivalent clamped-free viscoelastic layer. As the frequency increases, M_3 jumps to the S_0 mode of the free elastic plate. After intersecting S_0, M_3 veers towards the M_3' mode, and as the frequency increases further, M_3 intersects first M_3' and then A_0. M_3 does not intersect the bilayer mode M_4 (see inset), but veers and tends asymptotically to the M_4' mode. As has been shown for SH waves,[25] the bilayer modes do not intersect each other. However, the presence of intersections between the two families of asymptotic solutions leads to the mode repulsion phenomenon[33] (see, for instance, M_3 and M_4), which is particularly evident when the elastic plate is

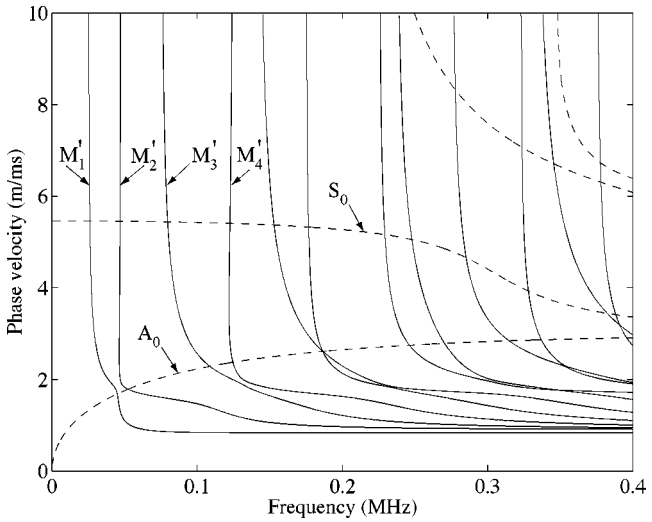


FIG. 4. Asymptotic solution for Lamb waves: (dashed lines) free metallic plate and (solid lines) clamped-free equivalent viscoelastic layer.

much stiffer and heavier than the viscoelastic layer.

Generally speaking, the bilayer cutoff frequencies tend to those of the clamped-free equivalent viscoelastic layer which correspond, for Lamb modes, to shear or longitudinal through-thickness standing waves and are given by

$$f_{\mathrm{S}}=\frac{c_{\mathrm{S}}^{v}}{4d_{v}}(2N-1), \tag{55}$$

$$f_{\mathrm{L}}=\frac{c_{\mathrm{L}}^{v}}{4d_{v}}(2N-1), \tag{56}$$

where $N\epsilon\{1,2,3,...\}$. Similarly, the cutoff frequencies of SH mode correspond to shear through-thickness standing waves in the equivalent viscoelastic layer and are also given by (55). As a consequence, all the bilayer cutoff frequencies of the SH modes occur around those of the Lamb modes and the first cutoff frequency is roughly the same in both cases.

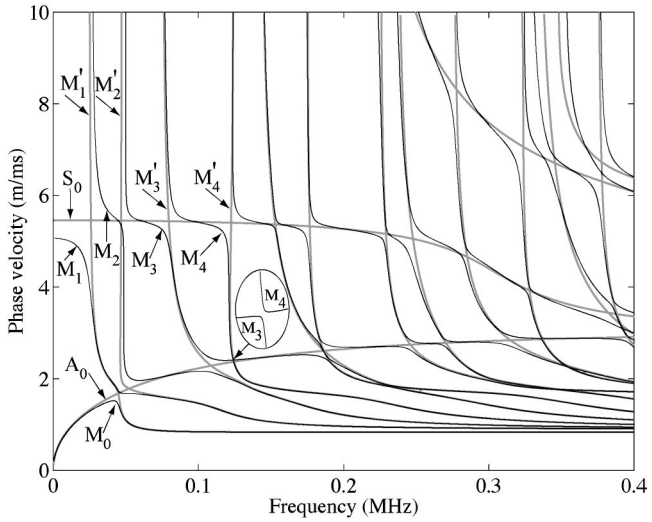


FIG. 5. Phase velocity dispersion curves: (gray lines) asymptotic solution and (black lines) bilayer modes.

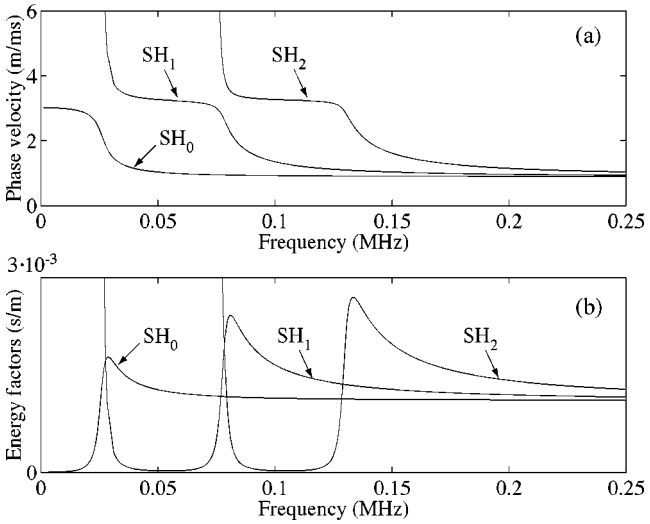


FIG. 6. Dispersion curves for SH modes propagating in the bilayer described in Table I: (a) phase velocity and (b) energy factor Q_E.

V. ELASTIC ENERGY FACTOR DISPERSION

In order to study the dispersion of the energy factors, the main properties of the dispersion of the Q_E factors are recalled first.[25] Figures 6(a) and (b) show the phase velocity and the Q_E factor for the first three SH modes of the bilayer (Table I). As the frequency tends to infinity, Q_E tends asymptotically to the reciprocal of the shear velocity of the equivalent viscoelastic layer

$$\lim_{f\to\infty} Q_E=\frac{1}{c_{\mathrm{S}}^{v}}. \tag{57}$$

Moreover, for all the modes whose order is higher than SH_0, Q_E minima occur at the Love transition frequencies, where the bilayer modes intersect the SH_0 mode of the free metallic plate. At these frequencies, the energy factor is independent of the mode order and is given by[25]

$$Q_{\mathrm{E}}=\frac{1}{c_{\mathrm{S}}^{v}}\frac{1}{c_{\mathrm{S}}^{v}/c_{\mathrm{S}}^{e}+2c_{\mathrm{S}}^{e}\rho_{e}d_{e}/c_{\mathrm{S}}^{v}\rho_{v}d_{v}}. \tag{58}$$

The frequency dependence of the energy factors Q_Γ and Q_Δ is much more complex. Figures 7(a) and (b) provide a comparison between the SH_0 mode and the M_1 mode of Fig. 3. While Q_E experiences only one peak [Fig. 7(b)] where the mode jumping occurs, Q_Γ and Q_Δ exhibit several peaks due to the more tortuous path of M_1. However, it can be noticed that as the frequency tends to infinity, the phase velocity of M_1 tends to that of SH_0. Moreover, Q_Γ tends to Q_E whereas Q_Δ vanishes. The reason for this behavior is that at high frequencies, the two SV bulk waves in the equivalent viscoelastic layer become dominant while the L and SV bulk waves in the metallic plate vanish.

Even though in the case shown in Fig. 7 Q_Δ is always lower than Q_Γ, this is not generally true for the other modes as will be shown later, and it is not possible to state general properties for all the modes as in the case of SH waves. On the other hand, for practical purposes, a mode can only be employed in the frequency ranges in which its phase velocity

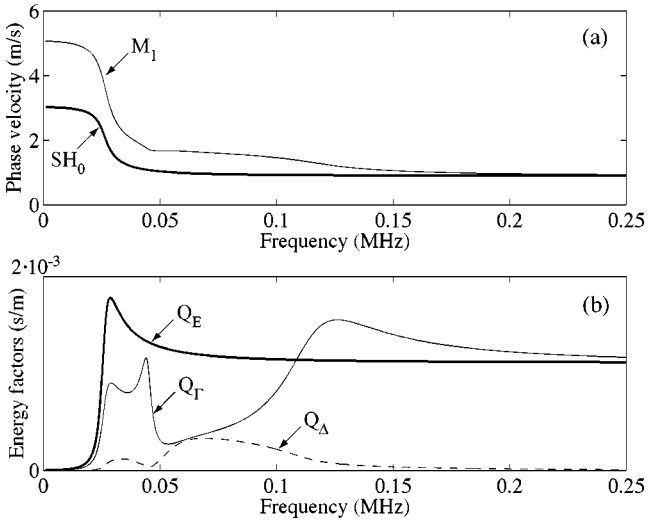


FIG. 7. (a) Phase velocity and (b) energy factor dispersion curves for the bilayer described in Table I: (solid bold line) Q_E for the SH_0 mode; (solid line) Q_Γ for the M_1 mode; and (dashed line) Q_Δ for the M_1 mode.

is not too dispersive since high dispersion is accompanied by a strong distortion and amplitude decay of the wave packet over long propagation distances. For this reason, in the following, the investigation of the bilayer properties is confined to the plateau regions of the dispersion curves. In Fig. 8 five different zones, labeled A, B, C, etc., have been identified. The corresponding energy factor curves are shown in Fig. 9.

In zone A, Q_E is always larger than Q_Γ while Q_Δ is negligible. This is proved in the Appendix.

Close to the lower limit of zone B (Fig. 9), the relationship between the energy factors is the same as in zone A. However, as the frequency increases the inequality changes, and at the upper limit of zone B it is inverted (i.e., $Q_E < Q_\Gamma < Q_\Delta$). This is due to the fact that zone B ends at the second cutoff frequency of the bilayer Lamb modes. This frequency is characterized by longitudinal through-thickness standing waves which have large ε_0 which implies that $\delta \gg \gamma$ [see (33) and (34)].

Close to the lower limit of zone C (Fig. 9), the energy

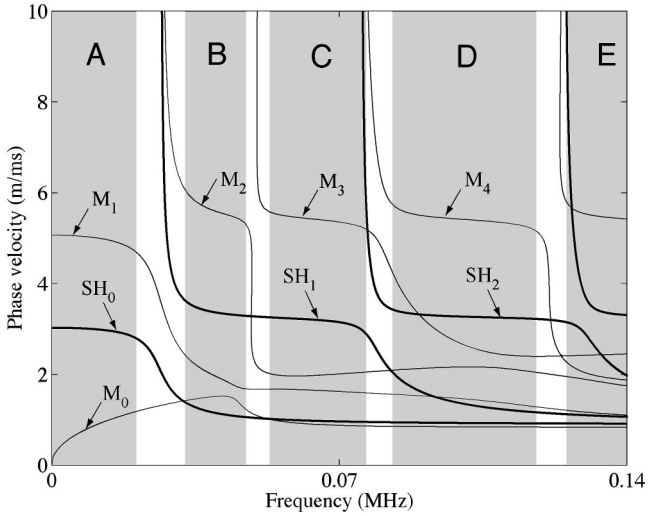


FIG. 8. Plateau zones in the dispersion curves of the bilayer: (solid lines) Lamb waves and (solid bold lines) SH waves.

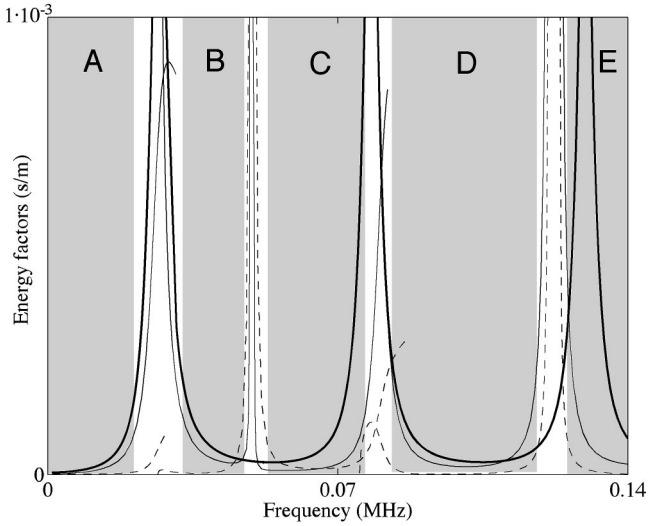


FIG. 9. Bilayer energy factors: (solid lines) Q_Γ; (dashed lines) Q_Δ; and (solid bold lines) Q_E.

factors behave as in the upper limit of zone B, while the situation is inverted as the frequency reaches the upper limit of zone C. Here, the third cutoff for Lamb and the second cutoff for SH waves occur. These cutoffs are characterized by shear through-thickness standing waves and consequently $Q_E > Q_\Gamma > Q_\Delta$. Similar considerations hold for all the other zones.

VI. LOW MATERIAL ATTENUATION

The analysis of the dispersion of the equivalent bilayer energy factors allows the guided wave attenuation to be obtained through (43) and (44). Figure 10 shows a comparison between the first-order approximation (dashed lines) and the numerical solution of the exact dispersion equation (gray lines), obtained by using the software Disperse.[24] The curves refer to the bilayer of Table I. In particular, the line labeled ζ_E refers to the guided wave attenuation of SH_0 for α_S

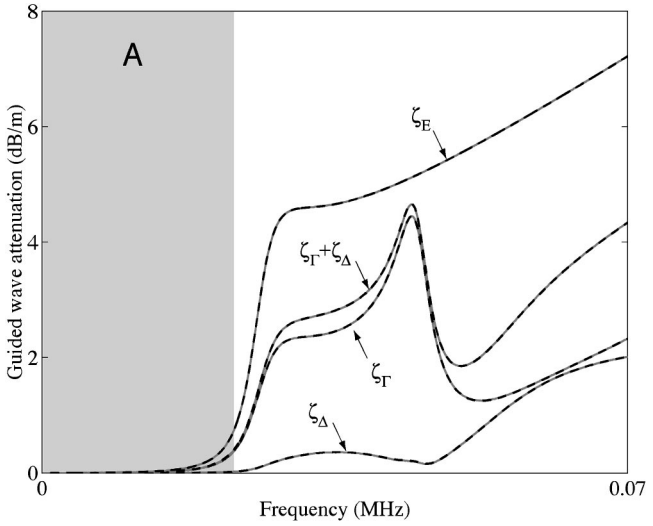


FIG. 10. Comparison between the first order approximation (dashed lines) and the numerical solution (solid gray lines): ζ_E attenuation of the SH_0 mode when $\alpha_S = 0.01$ np/wl; ζ_Δ attenuation of M_1 for $\alpha_S = 0$ and $\alpha_L = 0.01$ np/wl; ζ_Γ attenuation of M_1 for $\alpha_S = 0.01$ np/wl and $\alpha_L = 0$.

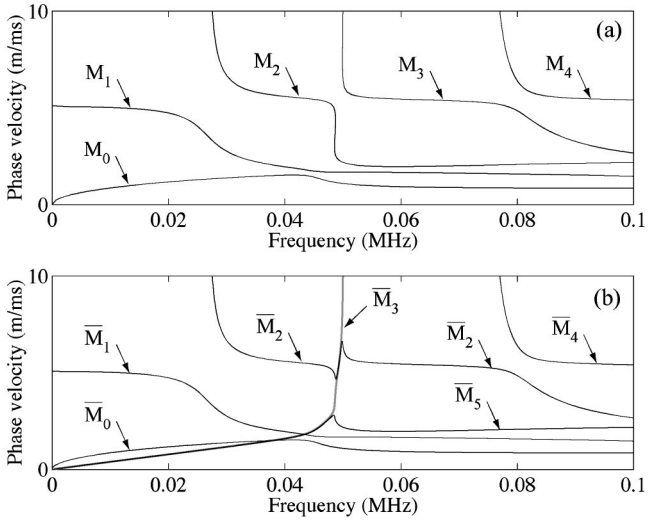


FIG. 11. Lamb wave phase velocity dispersion curves for the bilayer described in Table I: (a) elastic case; and (b) $\alpha_S=\alpha_L=0.01$ np/wl.

$=0.01$ np/wl. The curve ζ_Δ provides the attenuation of the Lamb mode M_1 for $\alpha_S=0$ np/wl and $\alpha_L=0.01$ np/wl. Conversely, ζ_Γ gives the guided wave attenuation associated with M_1 when $\alpha_S=0.01$ np/wl and $\alpha_L=0$ np/wl. Lastly $\zeta_\Gamma+\zeta_\Delta$ refers to the M_1 mode when $\alpha_S=\alpha_L=0.01$ np/wl. In all cases, the linear approximations, are in excellent agreement with the numerical solutions.

At low frequency (zone A), by considering that Q_Δ is lower than Q_Γ and noting that viscoelastic materials are characterized by $\alpha_L \ll \alpha_S$, it follows that the role of the longitudinal attenuation α_L is negligible compared to α_S. Also, as $Q_E > Q_\Gamma$ [see Eq. (A18)] the guided wave attenuation of the SH_0 mode is much larger than that of M_1.

As was discussed in Sec. III, in order to characterize the dispersion of the viscoelastic bilayer through the equivalent elastic bilayer, both the propagating and nonpropagating modes have to be considered. Figure 11(b) shows the phase velocity dispersion curves of Lamb waves for a slightly attenuative bilayer ($\alpha_S=\alpha_L=0.01$ np/wl). These curves almost overlap those of the elastic bilayer, shown in Fig. 11(a), with the exception of certain regions. In particular, the modes $\bar{M}_0$ and $\bar{M}_1$ are essentially the same as M_0 and M_1. As the frequency increases, the mode $\bar{M}_2$ follows M_2 up to 0.05 MHz and then suddenly joins M_3. The mode $\bar{M}_5$ overlaps M_2 above 0.05 MHz. However, at lower frequencies, $\bar{M}_5$ follows a path which does not belong to the dispersion curves of the elastic bilayer. Similar branches have been observed in the case of Lamb waves propagating in a free isotropic viscoelastic layer[34] and for an elastic plate loaded with an attenuative elastomer.[23] M_3 (shown in gray) overlaps $\bar{M}_5$ at low frequency, and then jumps to the trajectories of $\bar{M}_2$ and $\bar{M}_3$ [Fig. 11(a)]. Moreover, the power flow of $\bar{M}_3$ is negative as the energy flows along $-\hat{x}_1$. As a consequence, the attenuation and the group velocity are negative. The anomalous behavior of $\bar{M}_3$, and the presence of the unexpected branch of $\bar{M}_5$ at low frequency, which do not occur in the case of SH waves, are due to the presence of nonpropagating modes with non-purely-imaginary $\mathbf{k}_\parallel$ in the dispersion curves of the elastic bilayer.[35]

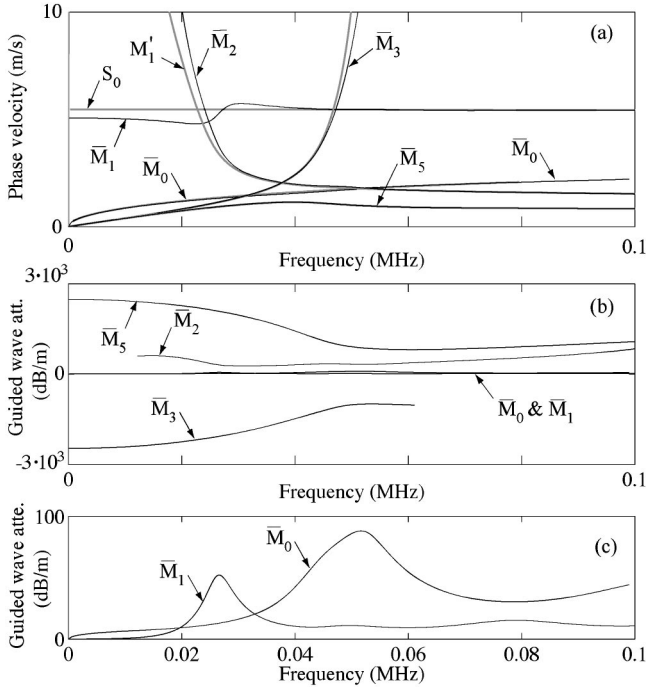


FIG. 12. Lamb wave dispersion curves for the bilayer described in Table I with $\alpha_S=\alpha_L=1$ np/wl: (gray lines) asymptotic solution; (black lines) bilayer modes: (a) phase velocity; (b) guided wave attenuation; and (c) attenuation of the $\bar{M}_0$ and $\bar{M}_1$ modes.

VII. LARGE MATERIAL DAMPING

High material damping is accompanied by large imaginary parts of the Lamé constants which make the zeroth- and first-order approximations less accurate. Moreover, the large dispersion of the Lamé moduli may cause high dispersion of the energy factors of the artificial bilayer (Sec. III). As a result, the energy factors obtained for the equivalent bilayer need to be corrected through the perturbation factors. However, as the frequency dependence of the acoustic properties of a viscoelastic material cannot be stated in a general form holding for all the different viscoelastic materials (whereas all Newtonian fluids can be described by one dispersion relationship which depends on one parameter, the viscosity, for a generic viscoelastic material an infinite number of parameters would be required), in the following, the perturbation factor will be neglected. Additionally, the bulk attenuation per wavelength is assumed not to vary with frequency.

Figure 12(a) shows the phase velocity dispersion curves for Lamb waves propagating in the bilayer described in Table I when a large material attenuation is considered. Such a bilayer is representative of steel plates coated with bitumen (see for bitumen properties Ref. 36). These curves are considerably different from those obtained in the elastic case [see Fig. 11(a)]. In particular, the mode jumping phenomenon, which occurs for the elastic bilayer, has now disappeared. For instance, the trajectory of the $\bar{M}_1$ mode oscillates around that of the free elastic plate mode S_0 rather than jumping to M_1', A_0 and M_2' as happened in the elastic bilayer (see Fig. 5). Similarly, the $\bar{M}_0$ mode follows the A_0 mode path rather than veering toward M_1'. On the other hand, $\bar{M}_2$ follows the trajectory of the clamped free viscoelastic layer M_1' mode rather than jumping to S_0. In gen-

eral, each bilayer mode follows the trajectory of one asymptotic mode only. A similar phenomenon occurs for Lamb wave propagation in free plastic plates. Chan *et al.*[37] showed that for a large material damping, the Lamb modes split into two families of modes which correspond to the modes of a fluid layer in vacuum (which has zero shear velocity) and those of an artificial layer in vacuum which only supports shear waves (with zero longitudinal velocity). These two families have the same role as the asymptotic modes discussed in this paper since Lamb modes in elastic plates can be thought of as an interaction between them as demonstrated by Überall *et al.*[33] A further example has been provided by Castaings *et al.*,[38] who considered the propagation of Lamb waves in a sandwich plate made of three anisotropic and viscoelastic layers. Also in this case, for large values of material absorption, modes no longer jump.

The absence of the mode jumping between the bilayer modes leads to the intersection of the phase velocity spectra of different modes. However, these intersections do not correspond to mode intersections, since two different modes would only intersect each other at a given frequency if the couples (C_{ph},ζ) were the same for both modes. As an example, M_1 and M_2 cross in phase velocity [Fig. 12(a)] but they do not in attenuation [Fig. 12(b)].

The $\overline{M}_3$ has negative attenuation. This is because the mode originates from the complex branches of the M_2 mode of the elastic bilayer whose power flow is directed along $-\hat{x}_1$.[35]

Figure 12(b) shows that the bilayer modes $\overline{M}_1$ and $\overline{M}_0$, which follow the S_0 and A_0 modes of the free elastic plate, have much lower attenuation than the modes that follow the clamped-free modes of the viscoelastic layer. This is due to the fact that in one case the energy is mainly trapped in the viscoelastic layer while, in the other, the energy flows primarily in the elastic plate. As a consequence, only the modes which tend to the free elastic plate modes can be practically employed. Figure 12(c) is a comparison between the attenuation spectra of $\overline{M}_0$ and $\overline{M}_1$. Since the attenuation of $\overline{M}_0$ is usually larger than that of $\overline{M}_1$ (especially at low frequency), $\overline{M}_1$ is more attractive than M_0 for traveling long distances in the bilayer. Figure 13 shows the dispersion curves of this mode (thick solid lines) when only a large bulk shear attenuation is considered. The first-order approximation is also shown [Fig. 13(b)]. It is interesting to note that the attenuation spectrum exhibits peaks which occur at the transitions where the phase velocity changes rapidly and mode jumping would occur if the bilayer was elastic. These peaks are due to a resonance phenomenon which causes maximum energy transfer from the metallic plate into the viscoelastic layer.[25] The resonance would not occur if the viscoelastic layer was infinitely thick.

The first-order approximation provides a satisfactory approximation in the plateau zones of the dispersion curves, while it becomes inaccurate around the guided wave attenuation maxima due to the rapid variation of the energy factors as shown in Fig. 13(b).

Figure 14 provides a comparison between $\overline{M}_1$ and SH_0. In this analysis the bulk longitudinal attenuation is neglected

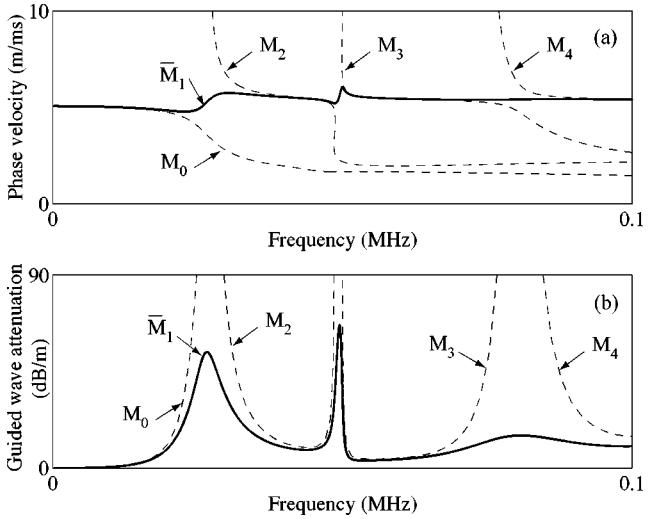


FIG. 13. Lamb dispersion curves for the bilayer described in Table I with $\alpha_S=1$ np/wl $\alpha_L=0$: (a) phase velocity for $\overline{M}_1$ (solid lines) and the first modes of the elastic bilayer (dashed lines); (b) attenuation of $\overline{M}_1$ (solid lines) and first-order approximation (dashed lines).

at first (solid lines). The two modes experience the first attenuation peak at the same frequency since the first cutoff frequency is the same for both modes [Eq. (55)]. Moreover, in the frequency interval around the first cutoff, the guided wave attenuation of the SH_0 mode is larger than that of $\overline{M}_1$. This follows from the fact that in the zone A (Fig. 9), SH_0 requires a higher strain energy than M_1 in order to produce unit power flow as has been shown in Sec. V. However, as the frequency approaches the second cutoff frequency of the Lamb mode M_3 [Eq. (56), upper limit of zone B in Fig. 9], $\overline{M}_1$ has another maximum, while the SH_0 attenuation tends to its minimum. On the other hand, if the bulk longitudinal attenuation is considered (dashed line), the second sharp peak disappears. At this frequency the strain energies, δ and γ [Eqs. (33) and (34)], are primarily carried by longitudinal partial waves. As a consequence, if the longitudinal attenuation increases, the viscoelastic layer behaves as if it were infinitely thick (since all the energy transmitted from the me-

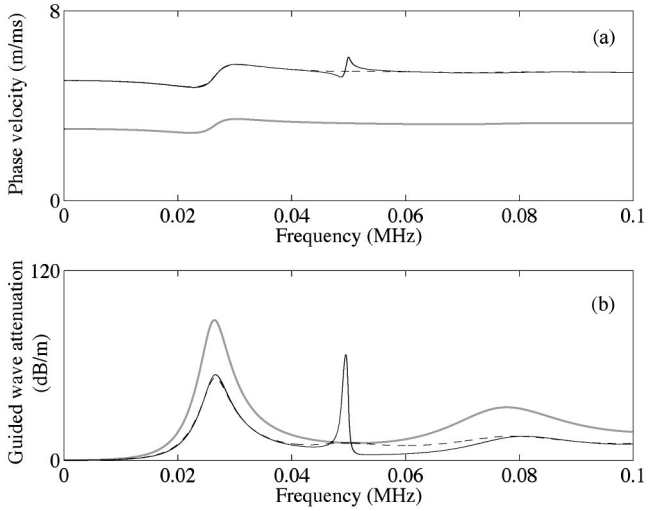


FIG. 14. Dispersion curves for the bilayer of Table I with $\alpha_S=1$ np/wl: (solid black line) SH_0; (solid gray line) $\overline{M}_1$ with $\alpha_L=0$; (dashed line) $\overline{M}_1$ with $\alpha_L=1$ np/wl.

tallic plate is lost into the bulk of the viscoelastic layer) and the peak vanishes. Moreover, the M_1 mode is sensitive to the longitudinal bulk attenuation around the peak frequencies given by (56) only since, over the rest of the spectrum, the energy in the viscoelastic layer travels in the form of γ energy rather than δ according to Sec. V.

VIII. CONCLUSIONS

The propagation of Lamb waves in metallic plates coated with viscoelastic layers has been investigated.

A general formulation relating the guided wave attenuation to the amount of strain energy stored in the viscoelastic layer per unit in-plane power flow has been derived. It has been shown that the strain energy can be split in two parts, δ and γ, which account for the effects of the two material bulk attenuations (longitudinal and shear) on the attenuation of the guided wave separately.

The topology of the dispersion curves of the elastic bilayer experiences substantial modifications when internal damping is considered. In particular, due to the energy dissipation within the attenuative layer, a mode is always propagating and its dispersion strongly depends on the magnitude of the internal damping. Two different material attenuation regimes can be defined. In the low regime, a bilayer mode embraces both the nonpropagating and propagating branches of the corresponding elastic mode. The non-purely-imaginary complex branches of the elastic bilayer dispersion curves cause anomalies in the phase velocity and guided wave attenuation spectra. For large values of the material attenuation, the modes can be classified as (a) modes whose energy travels primarily in the elastic plate and (b) modes whose energy is trapped in the viscoelastic layer. While the modes of the second family have little practical interest, as they are highly attenuated with distance, the modes of the first family can be employed in suitable frequency ranges. In particular, the attenuation spectra of these modes exhibit peaks which occur roughly at the through-thickness resonance frequencies of the viscoelastic layer if it were considered elastic and rigidly clamped at the bilayer interface.

The comparison between Lamb and SH waves suggests that the guided wave attenuation of SH waves is considerably larger than that of Lamb modes before and immediately after the first attenuation peak. At higher frequencies, the Lamb wave attenuation may be higher than that of the SH mode, depending on the magnitude of the longitudinal attenuation. However, apart from this region, the guided wave attenuation is not sensitive to the magnitude of the longitudinal bulk attenuation.

ACKNOWLEDGMENTS

The author is grateful to Professor P. Cawley and Dr. M. J. S. Lowe for critical reading of the manuscript and for many helpful discussions.

APPENDIX: ENERGY FACTORS AT LOW FREQUENCY

In order to show the validity of the inequality $Q_E > Q_\Gamma > Q_\Delta$ in zone A, let us consider the limit case as the frequency approaches zero. For both SH and Lamb waves, the displacement field tends to be constant across the thickness of the bilayer [i.e., $h_2(x_3)=\overline{h_2}$ and $h_1(x_3)=\overline{h_1}$ see Eqs. (16) and (18)]. For SH waves, the only nonzero strain is ε_{12} and the peak strain energy in the equivalent viscoelastic layer can be obtained by substituting the displacement expression (16) into ε_{12} and integrating (25) over the viscoelastic layer thickness:

$$\frac{dE}{dx}=\frac{1}{2}\tilde{\mu}_r^v d_v \mathbf{K}_\parallel^2 \overline{h_2^2}. \tag{A1}$$

The in-plane power flow is obtained by integrating the real part of the Poynting vector component along $\hat{x}_1$ over the thickness

$$\langle P\rangle=\tfrac{1}{2}\omega(\tilde{\mu}_r^e d_e+\tilde{\mu}_r^v d_v)\mathbf{K}_\parallel \overline{h_2^2}, \tag{A2}$$

thus, by the definition of Q_E in (29),

$$Q_E=\frac{1}{C_{ph}}\frac{\tilde{\mu}_r^v d_v}{\tilde{\mu}_r^v d_v+\tilde{\mu}_r^e d_e}, \tag{A3}$$

where the phase velocity can be obtained by considering the propagation of a plane wave along the bilayer

$$C_{ph}^2=\frac{\tilde{\mu}_r^v d_v+\tilde{\mu}_r^e d_e}{\rho_v d_v+\rho_e d_e}. \tag{A4}$$

If $\tilde{\mu}_r^v d_v/\tilde{\mu}_r^e d_e \ll 1$, the expression (A3) in the first-order approximation takes the form

$$Q_E\simeq\frac{1}{C_{ph}}\frac{\tilde{\mu}_r^v d_v}{\tilde{\mu}_r^e d_e}. \tag{A5}$$

In the case of Lamb waves the traction-free condition at the two boundaries of the bilayer implies that the stresses σ_{33} and σ_{31} are negligible compared to σ_{11} and σ_{22} across the thickness. Under this hypothesis, from the consitutive equations (9) the strain tensor components are

$$\varepsilon_{33}=-\frac{\tilde{\lambda}_r^v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\varepsilon_{11}, \tag{A6}$$

$$\varepsilon_{31}=0, \tag{A7}$$

in which (A6) takes into account the Poisson effect for plane strain deformation. The strain $\overline{\varepsilon_{11}}$ is derived from the displacement (18) with $h_1(x_3)=\overline{h_1}$. By substituting the strains (A6) and (A7) into (33) and (34) and integrating the expressions so obtained over the viscoelastic layer thickness, one obtains

$$\frac{d\Delta}{dx}=\tilde{\mu}_r^v\frac{2\tilde{\mu}_r^v d_v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\mathbf{K}_\parallel^2\overline{h_1^2}, \tag{A8}$$

$$\frac{d\Gamma}{dx}=\tilde{\lambda}_r^v\frac{2\tilde{\mu}_r^v d_v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\mathbf{K}_\parallel^2\overline{h_1^2}. \tag{A9}$$

The in-plane power flow is

$$\langle P\rangle=2\omega\left(\tilde{\mu}_r^e d_e\frac{\tilde{\lambda}_r^e+\tilde{\mu}_r^e}{\tilde{\lambda}_r^e+2\tilde{\mu}_r^e}+\tilde{\mu}_r^v d_v\frac{\tilde{\lambda}_r^v+\tilde{\mu}_r^v}{\tilde{\lambda}_r^v+2\tilde{\mu}_r^v}\right)\mathbf{K}_\parallel\overline{h_1^2}, \tag{A10}$$

and

$$Q_{\Delta}=\frac{1}{C_{ph}}\frac{\tilde{\mu}_r^v\tilde{\mu}_r^v d_v/(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v)}{(\tilde{\mu}_r^e d_e(\tilde{\lambda}_r^e+\tilde{\mu}_r^e)/(\tilde{\lambda}_r^e+2\tilde{\mu}_r^e)+\tilde{\mu}_r^v d_v(\tilde{\lambda}_r^v+\tilde{\mu}_r^v)/(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v))}, \tag{A11}$$

where the phase velocity is given by

$$C_{ph}^2=4\frac{(\tilde{\mu}_r^e d_e(\tilde{\lambda}_r^e+\tilde{\mu}_r^e)/(\tilde{\lambda}_r^e+2\tilde{\mu}_r^e)+\tilde{\mu}_r^v d_v(\tilde{\lambda}_r^v+\tilde{\mu}_r^v)/(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v))}{\rho_e d_e+\rho_v d_v}. \tag{A12}$$

Relationship (A11), in the first-order approximation, provides

$$Q_{\Delta}\simeq\tilde{\mu}_r^v\frac{1}{C_{ph}}\frac{\tilde{\mu}_r^v d_v}{\tilde{\mu}_r^e d_e}\frac{\tilde{\lambda}_r^e+2\tilde{\mu}_r^e}{(\tilde{\lambda}_r^e+\tilde{\mu}_r^e)(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v)}, \tag{A13}$$

and similarly

$$Q_{\Gamma}\simeq\tilde{\lambda}_r^v\frac{1}{C_{ph}}\frac{\tilde{\mu}_r^v d_v}{\tilde{\mu}_r^e d_e}\frac{\tilde{\lambda}_r^e+2\tilde{\mu}_r^e}{(\tilde{\lambda}_r^e+\tilde{\mu}_r^e)(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v)}. \tag{A14}$$

As a consequence, by taking the ratio of (A14) to (A13),

$$\frac{Q_{\Gamma}}{Q_{\Delta}}\simeq\frac{\tilde{\lambda}_r^v}{\tilde{\mu}_r^v}, \tag{A15}$$

which proves that, at low frequencies, $Q_{\Gamma}\gg Q_{\Delta}$. It is interesting to note that, despite in zone A for a Lamb mode the shear bulk waves are dominated by the longitudinal, Q_{Γ} is larger than Q_{Δ} as the γ energy depends on both longitudinal and shear bulk waves as explained in Sec. III.

The comparison between Q_{Γ} and Q_E leads to

$$\frac{Q_E}{Q_{\Gamma}}\simeq\frac{C_{ph}^{\mathrm{Lamb}}}{C_{ph}^{\mathrm{SH}}}\frac{(\tilde{\lambda}_r^e+\tilde{\mu}_r^e)(\tilde{\lambda}_r^v+2\tilde{\mu}_r^v)}{\tilde{\lambda}_r^v(\tilde{\lambda}_r^e+2\tilde{\mu}_r^e)}. \tag{A16}$$

Moreover, for $\tilde{\mu}_r^v d_v/\tilde{\mu}_r^e d_e\ll 1$

$$\frac{Q_E}{Q_{\Gamma}}\simeq 2\left(\frac{\tilde{\lambda}_r^e+\tilde{\mu}_r^e}{\tilde{\lambda}_r^e+2\tilde{\mu}_r^e}\right)^{3/2}\left(1+2\frac{\tilde{\mu}_r^v}{\tilde{\lambda}_r^v}\right), \tag{A17}$$

which, when the metal considered is steel, gives

$$\frac{Q_E}{Q_{\Gamma}}\simeq 1+2\frac{\tilde{\mu}_r^v}{\tilde{\lambda}_r^v}. \tag{A18}$$

It can be concluded that, in the zone A, $Q_E>Q_{\Gamma}>Q_{\Delta}$. Moreover, from (43) and (44) it follows that SH modes are more attenuated than Lamb waves.

[1] M. G. Silk and K. F. Bainton, "The propagation in metal tubing of ultrasonic wave modes equivalent to lamb waves," Ultrasonics **17**(1), 11–19 (1979).

[2] W. Bottger, H. Schneider, and W. Weingarten, "Prototype emat system for tube inspection with guided ultrasonic waves," Nucl. Eng. Des. **102**, 356–376 (1987).

[3] J. L. Rose, J. J. Ditri, A. Pilarski, K. Rajana, and F. Carr, "A guided wave inspection technique for nuclear steam generator tubing," NDT & E Int. **27**, 307–310 (1994).

[4] W. Mohr and P. Holler, "On inspection of thin-walled tubes for transverse and longitudinal flaws by guided ultrasonic waves," IEEE Trans. Sonics Ultrason. **SU-23**, 369–378 (1976).

[5] D. N. Alleyne and P. Cawley, "The interaction of lamb waves with defects," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **39**, 381–397 (1992).

[6] D. N. Alleyne and P. Cawley, "The excitation of lamb waves in pipes using dry coupled piezoelectric transducers," J. Nondestruct. Eval. **15**, 11–20 (1996).

[7] D. N. Alleyne and P. Cawley, "Long range propagation of lamb waves in chemical plant pipework," Mater. Eval. **55**, 504–508 (1997).

[8] D. N. Alleyne, M. J. S. Lowe, and P. Cawley, "The reflection of guided waves from circumferential notches in pipes," J. Appl. Mech. **65**, 635–641 (1998).

[9] M. J. S. Lowe, D. N. Alleyne, and P. Cawley, "The mode conversion of a guided wave by a part-circumferential notch in a pipe," J. Appl. Mech. **65**, 649–656 (1998).

[10] P. Cawley, M. J. S. Lowe, F. Simonetti, C. Chevalier, and A. G. Roosenbrand, "The variation of the reflection coefficient of extensional guided waves in pipes from defects as a function of defect depth, axial extent, circumferential extent and frequency," J. Mech. Eng. Sci. **216**(C), 1131–1143 (2002).

[11] P. Fromme, P. Wilcox, and P. Cawley, "Remote monitoring of plate-like structures using guided wave arrays," in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 2002), Vol. 22A, pp. 157–164.

[12] P. Wilcox, "Guided wave beam steering and modal selectivity from omnidirectional transducer arrays," in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 2002), Vol. 22A, pp. 761–768.

[13] O. Diligent, T. Grahn, A. Boström, P. Cawley, and M. Lowe, "The low-frequency reflection and scattering of the s_0 lamb mode from a circular through-thickness hole in a plate: Finite element, analytical and experimental studies," J. Acoust. Soc. Am. **112**, 2589–2601 (2002).

[14] A. Demma, P. Cawley, M. Lowe, and A. G. Roosenbrand, "The reflection of fundamental torsional mode from cracks and notches in pipes," J. Acoust. Soc. Am. **114**(3), 611–625 (2003).

[15] T. Kundu, "Scattering of torsional waves by a circular crack in a transversely isotropic solid," J. Acoust. Soc. Am. **88**, 1975–1980 (1990).

[16] D. N. Alleyne, B. Pavlakovic, M. J. S. Lowe, and P. Cawley, "Rapid, long range inspection of chemical plant pipework using guided waves," Insight **43**, 93–96,101 (2001).

[17] P. Cawley, M. J. S. Lowe, D. N. Alleyne, B. Pavlakovic, and P. Wilcox, "Practical long range guided wave testing: application to pipes and rail," Mater. Eval. **61**(1), 66–74 (2003).

[18] R. L. Weaver and Y. Pao, "Axisymmetric elastic waves excited by a point source in a plate," J. Appl. Mech. **49**, 821–836 (1982).

[19] J. P. Jones, "Wave propagation in a two-layered medium," J. Appl. Mech. **31**(2), 213–222 (1964).

[20] J. Laperre and W. Thys, "Experimental and theoretical study of lamb wave dispersion in aluminium/polymer bilayers," J. Acoust. Soc. Am. **94**, 268–278 (1993).

[21] Z. Zhu and J. Wu, "The propagation of lamb waves in a plate bordered with a viscous liquid," J. Acoust. Soc. Am. **98**, 1057–1064 (1995).

[22] A. H. Nayfeh and P. B. Nagy, "Excess attenuation of leaky lamb waves due to viscous fluid loading," J. Acoust. Soc. Am. **101**, 2649–2658 (1997).

[23] B. Drinkwater, M. Castaings, and B. Hosten, "The measurement of a_0 and s_0 lamb wave attenuation to determine the normal and shear stiffnesses of a compressively loaded interface," J. Acoust. Soc. Am. **113**, 3161–3170 (2003).

[24] B. N. Pavlakovic, M. J. S. Lowe, D. N. Alleyne, and P. Cawley, "Disperse: A general purpose program for creating dispersion curves," in *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti (Plenum, New York, 1997), Vol. 16, pp. 185–192.

[25] F. Simonetti and P. Cawley, ''On the nature of shear horizontal wave propagation in elastic plates coated with viscoelastic materials,'' Proc. R. Soc. in press (2004).
[26] R. M. Christensen, *Theory of Viscoelasticity: An Introduction* (Academic, New York, 1971).
[27] A. H. Nayfeh, ''The propagation of horizontally polarized waves in multilayered anisotropic media,'' J. Acoust. Soc. Am. **86**, 2000–2012 (1989).
[28] A. H. Nayfeh, *Wave Propagation in Layered Anisotropic Media with Application to Composites* (Elsevier, Amsterdam, 1995).
[29] B. A. Auld, *Acoustic Fields and Waves in Solids, volume 1* (Krieger, Malabar, FL, 1990).
[30] B. A. Auld, *Acoustic Fields and Waves in Solids, volume 2* (Krieger, Malabar, FL, 1990).
[31] C. L. Yapura and V. K. Kinra, ''Guided waves in a fluid-solid bilayer,'' Wave Motion **21**, 35–46 (1995).
[32] R. Vein and P. Dale, *Determinants and their Applications in Mathematical Physics* (Springer, New York, 1999).
[33] H. Überall, B. Hosten, M. Deschamps, and A. Gérard, ''Repulsion of phase-velocity dispersion curves and the nature of plate vibration,'' J. Acoust. Soc. Am. **92**, 908–917 (1994).
[34] A. Bernard, M. J. S. Lowe, and M. Deschamps, ''Guided waves energy velocity in absorbing and non-absorbing plates,'' J. Acoust. Soc. Am. **110**, 186–196 (2001).
[35] F. Simonetti and M. J. S. Lowe, ''On the meaning of lamb mode nonpropagating branches,'' submitted to Proc. R. Soc. (2004).
[36] F. Simonetti and P. Cawley, ''A guided wave technique for the characterisation of highly attenuative viscoelastic materials,'' J. Acoust. Soc. Am. **114**, 158–165 (2003).
[37] C. W. Chan and P. Cawley, ''Lamb waves in highly attenuative plastic plates,'' J. Acoust. Soc. Am. **104**, 874–881 (1998).
[38] M. Castaings and B. Hosten, ''Inversion of ultrasonic, plane-wave transmission data in composite plates to infer viscoelastic material properties,'' J. Acoust. Soc. Am. **113**, 2622–2634 (2003).

Laser-generated thermoelastic acoustic sources in anisotropic materials

David H. Hurley[a)]
Idaho National Engineering and Environmental Laboratory, P.O. Box 1625, Idaho Falls, Idaho 83415-2209

(Received 15 December 2003; revised 27 January 2004; accepted 4 February 2004)

An analytical model appropriate for thermoelastic generation of acoustic waves in anisotropic materials is presented for both plane and line sources. The interaction of acoustic waves produced by subsurface sources with the bounding surface is accounted for using a method of images. For the plane source case, analytical solutions are found that form an appropriate basis for an angular spectrum of plane waves. For the line source case and for specific crystal symmetries and source orientations, it is shown in the limit of strong optical absorption, a buried line source is equivalent to applying a shear stress dipole at the bounding surface. However, contrary to the isotropic case, the character and strength of the equivalent surface stress is a function of propagation direction. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1690080]

PACS numbers: 43.35.Sx [YHB] Pages: 2054–2058

I. INTRODUCTION

Since the early 1960s there has been considerable interest in generating acoustic waves using pulsed laser irradiation. The early experimental work in this field used low repetition rate (10 Hz) Q-switched lasers with nanosecond pulse lengths for acoustic generation (nanosecond laser acoustics).[1–3] A large part of the modeling effort for nanosecond laser acoustics has relied on replacing the laser source with an equivalent set of stress boundary conditions.[2,4] This equivalent stress distribution, which takes the form of a shear stress dipole for isotropic material, has proven to be extremely useful in describing many features of laser generated (thermoelastic regime) acoustic waveforms. For instance, the epicental waveform,[5,6] the surface acoustic waveform,[7] and the displacement directivity pattern[8,9] are all modeled accurately using a set of stress boundary conditions that are equivalent to a thermoelastic source in the limit of strong optical absorption.

In the mid-1980s a second type of pulsed laser acoustics emerged that used high repetition rate ($\sim$80 MHz) ultrafast laser pulses ($\sim$100 fs) to generate and detect acoustic pulses with pulse durations of a few picoseconds.[10] The temporal-resolution-afforded picosecond acoustics has enabled researchers to concentrate on relating the source characteristics to fundamental physical processes.[11–13] The modeling effort for picosecond acoustics typically involves 1-D models that exploit the simple experimental geometry.[14]

For nanosecond acoustics, work specific to laser sources in anisotropic materials has been reported by Every and Sachse.[15] In their work, data interpretation concentrated on wavefront analysis. By contrast, Mourad *et al.*,[16] used the Cagniard–deHoop[17] method to obtain numerically the solutions for laser excitation in an anisotropic half-space where it was assumed, in analogy to the isotropic case, that the laser source could be modeled as a shear stress dipole applied at the material surface. In a similar approach Hurley and Spicer[18] developed analytical solutions for thermoelastic generation in transversely isotropic materials by a laser line and laser point source.

For picosecond acoustics, work specific to elastically anisotropic materials has been primarily experimental in nature and typically involves acoustic propagation in high symmetry directions.[19–22] Recently, Hurley *et al.*[23] have reported generation of picosecond shear waves using an ultrafast pump probe method. This study involved laser generation of longitudinal waves in an isotropic aluminum film. A portion of the longitudinal wave was mode converted to a shear wave at the interface between the isotropic film and anisotropic substrate. The shear acoustic waves were detected using an off-axis polarization sensitive detection scheme.

In this article both plane (1-D problem geometry) and line (2-D problem geometry) thermoelastic sources, appropriate for picosecond and nanosecond laser acoustics in elastically anisotropic half-spaces, are examined in detail using a method of images to satisfy the stress boundary conditions. The method of images in general involves finding the solution to a particular boundary-value-problem by introducing a fictitious source (image source) in a boundless medium that together with the actual source satisfies the boundary conditions. This method, which is employed routinely for electrostatic problems, does not find the same popularity in dynamic elasticity since, for typical applications, satisfying both the shear stress and normal stress conditions at the boundary is not possible.[24] For the 1-D case considered in this manuscript, the method of images is applied in a very straightforward manner enabling an analytical solution for the displacements. In the 2-D case, a hybrid version of the method of images is used to obtain an equivalent set of stress boundary conditions for a limited set of crystal symmetries/line-source orientations.

II. THEORETICAL DEVELOPMENT

The following equations provide a general description of thermoelastic generation in anisotropic materials (strain rate coupling has been neglected):

a)Electronic mail: hurldh@inel.gov

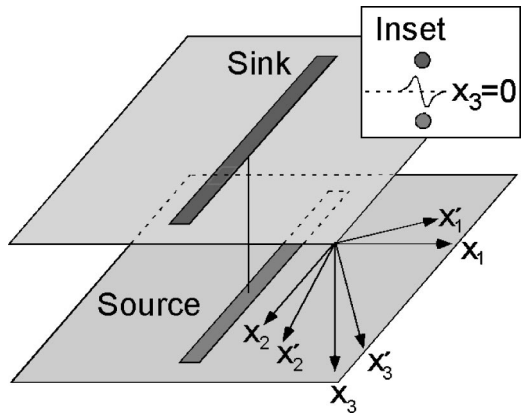


FIG. 1. Source/sink specifications. The axis/plane of the line/plane source/sink is perpendicular to the x_3 axis and parallel with the x_2 axis. The separation plane is perpendicular to the x_3 axis and is midway between the source and sink. The crystal axes coincide with the prime coordinate system. Inset: Conceptual illustration of the shear stress distribution at the separation plane resulting from the source/sink combination for the 2-D case.

$$\kappa'_{ij}T_{,ij}-\rho CT_{,t}+Q=0, \quad C'_{ijkl}u_{l,kj}-\rho u_{i,tt}=\beta'_{ij}T_{,j}. \tag{1}$$

The first equation is the linearized heat conduction equation and the second is the equation of motion for a linear elastic material. The material constants, κ'_{ij}, β'_{ij}, C'_{ijkl}, ρ, and C, are defined as the thermal conductivity tensor, the thermal pressure tensor, the elastic stiffness tensor, the mass density, and the specific heat respectively. The components of displacement and temperature are defined as u_i and T. In this development the crystal axes, which coincide with the prime coordinate system, have an arbitrary orientation relative to the unprimed axes (Fig. 1). The crystalline symmetry as it relates to the various tensor symmetries will be discussed in detail in the following sections. Since strain-rate coupling has been neglected, the solution to the heat conduction equations can be sought without considering the elastic equations of motion. Given that the spirit of this article is to provide analytical insight into the role of elastic and thermal anisotropy in thermoelastic generation of acoustic waves, heat diffusion will be neglected.[4,25] The temperature fields for both the 1-D and 2-D case are found by taking the temporal integral of the absorbed laser energy (q_0):

$$\begin{aligned} T_{1\text{-}D}&=T_0H(t)[\,\delta(x_3-a)-\delta(x_3+a)\,],\\ T_{2\text{-}D}&=T_0H(t)\,\delta(x_1)[\,\delta(x_3-a)-\delta(x_3+a)\,], \end{aligned} \tag{2}$$

where $T_0=q_0/\rho C$. The source, which is concentrated a distance $+a$ from the surface, is associated with an increase in temperature and the image source (sink), which is concentrated a distance $-a$ from the surface, is associated with a decrease in temperature (Fig. 1).[26] The solution procedure for the elastic displacements for the 1-D and 2-D cases involves applying a Fourier/Laplace transform to remove dependence on the spatial and temporal variables:

$$\begin{aligned} F(\eta,\kappa,s)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty}f(x_1,x_3,t)\\ \times e^{(i\eta x_1+i\kappa x_3-st)}dtdx_1dx_3. \end{aligned} \tag{3}$$

For the 1-D case the solution depends only on x_3 while the 2-D case depends on both x_1 and x_3. The transformed equations take the form

$$N\cdot\overline{u}=S, \tag{4}$$

where barred quantities refer to transformed variables and N and S which are functions of the transform parameters are define in the Appendix. The transformed displacements are decoupled algebraically giving

$$\begin{aligned} &\overline{u_i}=\frac{D_i}{D_0}, \quad D_1=\det(N_1),\\ &D_0=\det(N)=d_0(\kappa^2-\kappa_1^2)(\kappa^2-\kappa_2^2)(\kappa^2-\kappa_3^2), \end{aligned} \tag{5}$$

where N_i is obtained from N by replacing the ith column with S and κ_i are the roots of D_0. Using partial fractions the transformed displacements may be represented as

$$\begin{aligned} &\overline{u_i}=\sum_{j=1}^{3}\frac{A_{ij}}{d_0(\kappa^2-\kappa_j^2)}, \quad A_{ij}=\frac{D_i(\kappa_j)}{(\kappa_j-\kappa_{r1})(\kappa_j-\kappa_{r2})},\\ &r_1\neq j, \;\; r_2\neq j, \;\; r_2\neq r_1, \end{aligned} \tag{6}$$

where the amplitude coefficients, A_{ij}, are functions of propagation direction for elastically anisotropic materials.

A. One-dimensional case

The Fourier inversion now may be written as

$$\begin{aligned} &\overline{u_i}(x_3,s)=\frac{-\overline{A}_{ij}}{d_0}\sqrt{\frac{2}{\pi}}\frac{\partial}{\partial x_3}\left[\,\mathrm{Re}\int_0^{\infty}\frac{E(\omega,s,x_3)}{s^2(\omega^2-\omega_j^2)}d\omega\right],\\ &\overline{A}_{ij}=\frac{A_{ij}}{i\omega E(\omega,s,x_3)},\\ &E(\omega,s,x_3)=(e^{-i\omega s|x_3-a|}-e^{-i\omega s|x_3+a|}), \end{aligned} \tag{7}$$

where the substitution, $\kappa=s\omega$, has been made to facilitate the transform inversion.[27] The real ω axis is identified as the Fourier inversion path and the imaginary ω axis is identified as the Cagniard path.[17,28,29] Since all the poles lie on the imaginary ω axis, the Fourier inversion is accomplished by summing the residues along the imaginary ω axis:

$$\begin{aligned} \int_0^{\infty}\frac{E(\omega,s,x_3)}{(\omega^2-\omega_j^2)}d\omega&=\pi i\,\mathrm{Res}\left\{\frac{E(\omega,s,x_3)}{(\omega^2-\omega_j^2)}\right\}\\ &=\cdots\pi i\left(\frac{E(\omega_j,s,x_3)}{2\omega_j}\right). \end{aligned} \tag{8}$$

Given that the dependence on x_3 has been recovered, the stresses at the separation plane ($x_3=0$) can be evaluated:

$$\sigma_{3j}=C'_{3j13}u_{1,3}+C'_{3j23}u_{2,3}+C'_{3j33}u_{3,3}=0. \tag{9}$$

A detailed analysis of the stress can be circumvented by noting that the displacements are of the form

$$u_i\propto\frac{\partial^2}{\partial x_3^2}(e^{-\overline{\omega}s|x_3-a|}-e^{-\overline{\omega}s|x_3+a|}). \tag{10}$$

Since the displacement components are even functions about the plane $x_3=0$, it follows that the stresses at the separation

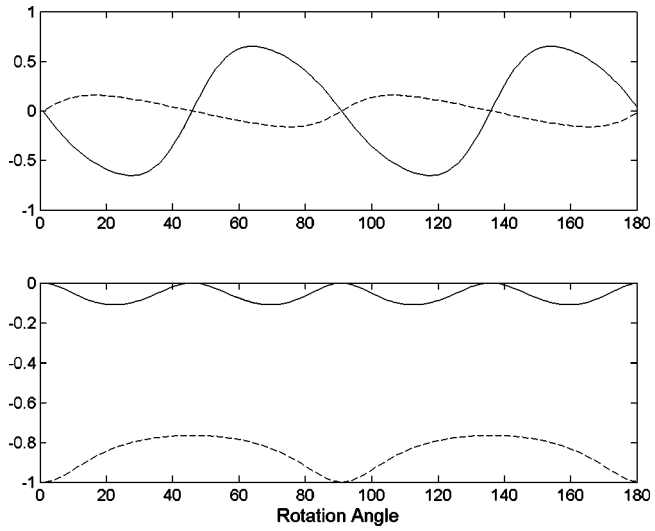


FIG. 2. Amplitude coefficients versus rotation about the x_2 axis [values have been normalized using $\overline{A}_{33}(\theta=0)$]. Top: solid line corresponds to normalized $\overline{A}_{11}$ [i.e., $G_{11}/G_{33}(\theta=0)$] and bottom line corresponds to normalized $\overline{A}_{13}$ [i.e., $G_{11}/G_{33}(\theta=0)$]. Bottom: solid and dashed line correspond to normalized $\overline{A}_{31}$ and $\overline{A}_{33}$, respectively.

plane must be zero. Thus for the 1-D case, a source and image sink satisfy stress free conditions at $x_3=0$ and, hence, the separation plane can be identified with the bounding plane of an elastic half-space. Furthermore, this combination of thermoelastic source and image sink in an unbounded medium is equivalent to a thermoelastic source in an elastic half-space. Since the image sink produces a modal amplitude distribution that is of equal magnitude and opposite sign to that of the source, mode converted waves are not required to satisfy the stress-free boundary conditions. This, however, is not the case for the somewhat similar problem of two half-spaces having different orientations.[23,30] Dependence on time is recovered using a table of Laplace transform-pairs giving

$$u_i(x_3,t)=\sqrt{\frac{\pi\overline{A}_{ij}^2}{2d_0^2}}\begin{Bmatrix} H(x_3-a)H(t-\overline{\omega}_j(x_3+a))\cdots \\ -H(a-x_3)H(t+\overline{\omega}_j(x_3+a)) \\ -H(x_3+a)H(t-\overline{\omega}_j(x_3-a))\cdots \\ +H(-x_3-a)H(t+\overline{\omega}_j(x_3-a)) \end{Bmatrix}. \tag{11}$$

As a check on the solution, it can be shown that the eigenvectors corresponding to the homogeneous solution can be formed from appropriate ratios of amplitude coefficients, A_{ij}. However, contrary to the homogeneous solution, the displacement waveforms given in Eqs. (11) form an appropriate bases for an angular spectrum of planes waves which could be used in the calculation of acoustic diffraction resulting from 2-D and 3-D thermal sources.[31]

To elucidate the effect of elastic anisotropy on the source strength, consider a material with cubic symmetry for which the amplitude coefficient can be separated into two terms:

$$\overline{A}_{ij}=T_0\beta_0 G_{ij}, \tag{12}$$

where β_0 is a function of both elastic and thermal constants[32] and G_{ij}, which is a function of only the elastic constants, gives the variation of the displacement amplitudes with propagation direction. The strong dependence of G_{ij} on propagation direction is illustrated in Fig. 2. The material constants are for single crystal Cu (cubic) and the initial orientation corresponds to the prime axes coinciding with the unprime axes. The solution is solved repetitively as the sample is rotated about the x_2 axis.

B. Two-dimensional case

For a two-dimensional source, analytical solutions exist only for certain crystal symmetries and for certain source/observation orientations.[7,18] However, instead of finding expressions for the displacements, the emphasis of this section will be on prescribing a set of stress boundary conditions that is equivalent to a center of thermal expansion located at the bounding surface.[33] Much of the groundwork has already been done in the previous section. The symmetries/orientations that will be discussed are given in Table I and correspond to cases for which D_0 can be split into two parts, one of second order and one of forth order in κ. Physically this corresponds to waves that are polarized in the (x_1,x_3) plane. Performing the same mathematical manipulations given in the preceding section, the displacements can be represented as

$$\begin{aligned} \overline{u}_1(\eta,\kappa,s)&=\frac{-i\eta\overline{A}_{1j}}{d_0}\frac{E(\kappa)}{\kappa^2-\kappa_j^2}, \\ \overline{u}_3(\eta,\kappa,s)&=\frac{-i\kappa\overline{A}_{3j}}{d_0}\frac{E(\kappa)}{\kappa^2-\kappa_j^2}, \\ E(\kappa)&=e^{ia\kappa}-e^{-ia\kappa}. \end{aligned} \tag{13}$$

Dependence on x_3 is regained using a table of transform pairs and is given as

$$\begin{aligned} \overline{u}_1(\eta,x_3,s)&=-\sqrt{\frac{\pi}{2}}\frac{i\eta\overline{A}_{1j}}{d_0\kappa_j}E(\kappa_j,x_3), \\ \overline{u}_3(\eta,x_3,s)&=\sqrt{\frac{\pi}{2}}\frac{\overline{A}_{3j}}{d_0\kappa_j}\frac{\partial E(\kappa_j,x_3)}{\partial x_3}, \\ E(\kappa_j)&=e^{-\kappa_j|x_3-a|}-e^{-\kappa_j|x_3+a|}. \end{aligned} \tag{14}$$

The next step in the process is to evaluate the stresses at the separation plane. Applying both a Laplace and Fourier transform to eliminate dependence on x_1 and time, the transformed stresses at the separation plane for all the cases excluding case 4 in Table I are represented as

$$\begin{aligned} \overline{\sigma}_{13}(\eta,x_3,s)&=-i\eta C'_{55}\overline{u}_3+C'_{55}\overline{u}_{1,3}, \\ \overline{\sigma}_{33}(\eta,x_3,s)&=-i\eta C'_{13}\overline{u}_1+C'_{33}\overline{u}_{3,3}. \end{aligned} \tag{15}$$

By noting that $\overline{u}_1$ and $\overline{u}_{3,3}$ are both odd functions of x_3, the normal stress evaluated at the separation plane is necessarily zero. However, the shear stress, σ_{13}, is nonzero at the separation plane. If, in addition to the source/sink combination, a shear stress of opposite sign and equal magnitude is applied at the separation plane, a stress free state is achieved at the separation plane (Fig. 1 inset). Thus the combination of (1) source, (2) sink and (3) shear stress applied at the

TABLE I. The seven crystal symmetries/line-source orientations considered. Crystalographic axes are defined as X, Y, Z and $R45°(x_1)$ refers to a 45° rotation about the x_1 axis. Cases 1–3 and 5–7 give the strength of the shear stress dipole, Γ_{13}, in terms of the elastic constants in the nonrotated frame (i.e., the crystallographic axes coincide with the unprimed axes).

Structure	Cubic (1)	Cubic (2)	Cubic (3a,3b)	Cubic (4)
Orientation	$R45°(x_1)$ $x_1 \| X$	$R45°(x_2)$ $x_2 \| Y$	(a) $R45°(x_3)$ $x_3 \| Z$ (b) $x_1, x_2, x_3 \| XYZ$	$R\theta(x_2), x_1 \| X, x_2 \| Y$
Γ_{13}/β_0	$\frac{C_{11}+2C_{44}-C_{12}}{C_{11}+C_{12}+2C_{44}}$	$\frac{4C_{44}}{C_{11}+C_{12}+2C_{44}}$	$\frac{C_{11}-C_{12}}{C_{11}}$	N/A
Structure	Hexagonal (5)	Tetragonal (6)	Orthorhombic (7)	
Orientation	$x_3 \| Z$	$x_1, x_2, x_3 \| XYZ$	$x_1, x_2, x_3 \| XYZ$	
Γ_{13}	$\frac{\beta_{11}C_{33}-\beta_{33}C_{13}}{C_{33}}$	$\frac{\beta_{11}C_{33}-\beta_{33}C_{13}}{C_{33}}$	$\frac{\beta_{11}C_{33}-\beta_{33}C_{13}}{C_{33}}$	

separation plane is equivalent to a source in an elastic half-space. At this point inverting the shear stress would be difficult if not impossible analytically. However, if the source and sink are brought together at the separation plane, they necessarily annihilate, leaving a very simple expression for the applied shear stress:

$$\overline{\sigma}_{13}(\eta, x_3=0, s) = \frac{-i\eta}{s\sqrt{2\pi}} T_0 \Gamma_{13},$$

$$\sigma_{13}(x_1, x_3=0, s) = T_0 \Gamma_{13} H(t) \frac{\partial \delta(x_1)}{\partial x_1}. \tag{16}$$

The functional form of the equivalent surface shear stress is a dipole, similar to the expression given by Rose.[4] However, in this case the amplitude of the dipole is a function of crystal symmetry and sample orientation. Table I gives the form of Γ_{13} for some symmetry/orientation combinations for which D_0 is separable in the fashion stated earlier.

Similar to the 1-D case, consider a cubic system for which Γ_{13} can be separated into two components, one that is purely elastic and one that depends on both thermal and elastic constants (cases 1–4 in Table I). Now consider cases 3a and 3b in detail. The surface normal for these two sample orientations coincides and, hence, experiments corresponding to 3a and 3b can be performed using the same sample. The amplitude of the shear stress dipole for cases 3a and 3b are identical. However, for a sample orientation between case 3a and 3b, the equivalent surface stress is no longer in the form of a shear stress dipole.

Now consider case 4 where the sample has been rotated by some arbitrary angle about the x_2 axis. For case 4, the transformed stresses take the from

$$\begin{aligned} \overline{\sigma}_{13}(\eta, x_3, s) &= -i\eta(C'_{15}\overline{u}_1 + C'_{55}\overline{u}_3) + C'_{55}\overline{u}_{1,3} \\ &\quad + C'_{35}\overline{u}_{3,3}, \\ \overline{\sigma}_{33}(\eta, x_3, s) &= -i\eta(C'_{13}\overline{u}_1 + C'_{35}\overline{u}_3) + C'_{35}\overline{u}_{1,3} \\ &\quad + C'_{33}\overline{u}_{3,3}. \end{aligned} \tag{17}$$

First note that in the rotated coordinate system, C'_{15} and C'_{35} are no longer zero and, as a result, the normal stress does not vanish. Second, the form of the transformed stress does not afford a simple analytical solution contrary to the other cases considered.

III. CONCLUSION

This paper examines the implication of elastic anisotropy as it relates to thermoelastic sources. For the 1-D case, a set of analytical solutions was found for all crystal symmetries/source orientations. These solutions form an appropriate bases set for an angular spectrum of plane waves. For the 2-D case in the limit of strong optical absorption and for specific crystal symmetries and source orientations, a buried line source was shown to be equivalent to applying a shear stress dipole at the bounding surface. However, contrary to the isotropic case, for an arbitrary sample orientation, the equivalent surface normal stress was nonzero and the equivalent surface shear stress was not a simple dipole. This has important implications since the temporal evolution of displacement waveforms is strongly related to the character of the source. For instance, in isotropic materials, displacement waveforms related to laser ablation,[9] which are accurately modeled using a normal point force, are markedly different than waveforms related to thermoelastic generation.[2,9]

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Dr. J. B. Spicer for helpful discussions during the course of this work. This work was sponsored by the U.S. Department of Energy, Office of Science-BES, Materials and Engineering Physics program under DOE Idaho Operations Office Contract No. DE-AC07-99ID13727.

APPENDIX

For the 1D case, N and S are defined as

$$N = \begin{bmatrix} C'_{55}\kappa^2 + \rho s^2 & C'_{45}\kappa^2 & C'_{35}\kappa^2 \\ C'_{45}\kappa^2 & C'_{44}\kappa^2 + \rho s^2 & C'_{34}\kappa^2 \\ C'_{35}\kappa^2 & C'_{34}\kappa^2 & C'_{33}\kappa^2 + \rho s^2 \end{bmatrix},$$

$$S = \begin{bmatrix} i\eta T_0 \beta'_{11}/\sqrt{2\pi}s \\ i\eta T_0 \beta'_{21}/\sqrt{2\pi}s \\ i\eta T_0 \beta'_{31}/\sqrt{2\pi}s \end{bmatrix}. \tag{A1}$$

Since C'_{ij} and β'_{ij} refer to a rotated coordinate system, the set of equations represented by Eq. (6) are valid for all crystal symmetries. In order to elucidate the effect of elastic anisotropy on the source strength, cubic symmetry will be assumed ($\beta_{ij}=\beta_0\delta_{ij}$). This allows the amplitude coefficients to be factored into two terms, one containing both elastic and thermal terms and one containing elastic terms only:

$$\overline{A}_{1j}/T_0\beta_0=\frac{\omega_j^2(C'_{35}C'_{44}\omega_j^2-C'_{45}C'_{34}\omega_j^2-C'_{35}\rho)}{(\kappa_j-\kappa_{r1})(\kappa_j-\kappa_{r2})},$$

$$\overline{A}_{2j}/T_0\beta_0=\frac{\omega_j^2(C'_{34}C'_{55}\omega_j^2-C'_{45}C'_{35}\omega_j^2-C'_{34}\rho)}{(\kappa_j-\kappa_{r1})(\kappa_j-\kappa_{r2})},$$

$$\overline{A}_{3j}/T_0\beta_0=\frac{\omega_j^4(C'^2_{45}-C'_{55}C'_{44})+\omega_j^2\rho(C'_{55}+C'_{44})-\rho^2}{(\kappa_j-\kappa_{r1})(\kappa_j-\kappa_{r2})}, \quad \text{(A2)}$$

$$r_1\neq j,\ \ r_2\neq j,\ \ r_2\neq r_1.$$

For the sake of brevity, N and S for the 2-D problem will be given for case 3 of Table I:

$$N_{11}=C_{11}\eta^2+C_{44}\kappa^2+\rho s^2,\ \ N_{13}=(C_{44}+C_{12})\eta\kappa=N_{31},$$
$$N_{22}=(\eta^2+\kappa^2)C_{44}+\rho s^2,\ \ N_{33}=C_{44}\eta^2+C_{11}\kappa^2+\rho s^2, \quad \text{(A3)}$$
$$N_{12}=N_{21}=N_{23}=N_{32}=0,\ \ S=\begin{bmatrix} i\eta T_0\beta_0/2\pi s \\ 0 \\ i\kappa T_0\beta_0/2\pi s \end{bmatrix}.$$

[1] R. M. White, "Excitation of Surface Elastic Waves by Transient surface Heating," Appl. Phys. Lett. **12**, 12–15 (1963).

[2] C. B. Scruby, R. J. Dewhurst, D. A. Hutchins, and S. B. Palmer, "Quantitative studies of thermally generated elastic waves in laser-irradiated metals," J. Appl. Phys. **51**, 6210–6216 (1980).

[3] D. A. Hutchins and A. C. Tam, "Pulsed photoacoustic materials characterization," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **33**, 439–449 (1986).

[4] L. R. F. Rose, "Point-source representation for laser-generation ultrasound," J. Acoust. Soc. Am. **75**, 723 (1984).

[5] Finite optical penetration and thermal diffusion lead to a precursor spike in the epicentral waveform which can not be accounted for using a stress dipole.

[6] K. L. Telschow and R. J. Conant, "Optical and thermal parameter effects on laser-generated ultrasound," J. Acoust. Soc. Am. **88**, 1494–1502 (1990).

[7] D. H. Hurley, J. B. Spicer, J. W. Wagner, and T. W. Murray, "Investigation of the anisotropic nature of laser-generated ultrasound in zinc and unidirectional carbon epoxy composites," Ultrasonics **36**, 355–360 (1998).

[8] J. R. Bernstein and J. B. Spicer, "Line source representation for laser-generated ultrasound in aluminum," J. Acoust. Soc. Am. **107**, 1352–1357 (2000).

[9] S. J. Davies, C. Edwards, G. S. Taylor, and S. B. Palmer, "Laser-generated ultrasound: its properties, mechanisms and multifarious applications," J. Phys. D **26**, 329–348 (1993).

[10] C. Thomsen, H. T. Grahn, H. J. Maris, and J. Tauc, "Picosecond interferometric technique for study of phonons in the Brillouin frequency range," Opt. Commun. **60**, 55–58 (1986).

[11] G. Tas and H. J. Maris, "Electron diffusion in metals studied by picosecond ultrasonics," Phys. Rev. B **49**, 15046–15054 (1994).

[12] T. Q. Qiu and C. L. Tien, "Short-pulse laser heating on metals," J. Heat Transfer **35**, 719–726 (1992).

[13] V. E. Gusev and O. B. Wright, "Ultrafast nonequilibrium dynamics of electrons in metals," Phys. Rev. B **57**, 2878–2888 (1998).

[14] Typically the lateral dimension of the laser source is much larger than the thickness of the film being investigated.

[15] A. G. Every and W. Sachse, "Imaging of Laser-Generated Ultrasonic Waves in Silicon," in *Acoustical Imaging, Vol. 19*, edited by H. Emert and H.-P. Harjes, (Plenum, New York, 1992), pp. 743–748.

[16] A. Mourad, M. Deschamps, and B. Castagnede, "Acoustic waves generated by a transient line source in an anisotropic half-space," Acustica **82**, 839–851 (1996).

[17] L. Cagniard, *Reflection and Refraction of Prog. Seismic Waves*, (McGraw–Hill, New York, 1962); A. T. de Hoop, Appl. Sci. Res., Sect. B **8**, 349 (1960).

[18] D. H. Hurley and J. B. Spicer, "Point-source representation for laser-generated ultrasound in an elastic, transversely isotropic half-space," J. Appl. Phys. **86**, 3423–3427 (1999).

[19] W. Chen, H. J. Maris, Z. R. Wasilewski, and S. Tamura, "Attenuation and velocity of 56 GHz longitudinal phonons in gallium arsenide from 50 to 300 K," Philos. Mag. B **70**, 687–698 (1994).

[20] O. Matsuda, I. Ishii, T. Fukui, J. J. Baumberg, and O. B. Wright, "Wavelength selective photoexcitation of picosecond acoustic-phonon pulses in a triple GaAs/$Al_{0.3}Ga_{0.7}$As quantum well structure," Physica B **316–317**, 205–208 (2002).

[21] O. B. Wright, B. Perrin, O. Matsuda, and V. E. Gusev, "Ultrafast carrier diffusion in gallium arsenide probed with picosecond acoustic pulses," Phys. Rev. B **64**, 81202-81205 (2001).

[22] The laser source corresponding to Refs. 19–21 was nonthermal.

[23] D. H. Hurley, O. B. Wright, O. Matsuda, V. E. Gusev, and O. V. Kolosov, "Laser picosecond acoustics in isotropic and anisotropic materials," Ultrasonics **38**, 470–474 (2000).

[24] A physical interpretation of this difficulty is easily realized for the case of a source/sink combination in an elastically isotropic material. In this case the image sink and source only produce longitudinal waves and, hence, mode conversion from longitudinal to shear waves at the interface cannot be adequately described with a source and image sink.

[25] J. F. Ready, *Effects of High-Powered Laser Radiation* (Academic, New York, 1971).

[26] The current solution serves as a Green's function which can be used to find the solution due to a distributed source.

[27] It should be noted that by relating $i\omega s$ with the partial derivative with respect to x_1, the remaining integrand is even in ω and hence the integration limits can be changed to the positive ω axis.

[28] R. G. Payton, *Elastic Wave Propagation in Transversely Isotropic Media* (Martinus Nijhoff, The Hague, 1983).

[29] D. H. Hurley, Ph.D. dissertation, Johns Hopkins University, 1997.

[30] B. A. Auld, *Acoustic Fields and Waves in Solids* (Wiley-Interscience, New York, 1973), Vol. 2, pp. 38–43.

[31] M. S. Kharusi and G. W. Farnell, "Plane ultrasonic transducer diffraction fields in highly anisotropic crystals," J. Acoust. Soc. Am. **48**, 665–670 (1970).

[32] J. A. Hildebrand, "Thermoacoustic generation in anisotropic media," J. Acoust. Soc. Am. **79**, 1457–1460 (1985).

[33] In order to correctly account for mode conversion (3-D problem geometry) due to non-normal reflections in isotropic materials, Rose[4] considered a thermoelastic point-source to be buried some distance below the free surface. For an anisotropic half-space, the procedure, presented by Rose, is long and cumbersome owing to the fact that the shear and longitudinal deformations can not be decoupled.

Thermoacoustic enrichment of the isotopes of neon

D. A. Geller and G. W. Swift
Condensed Matter and Thermal Physics Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 30 October 2003; accepted for publication 28 January 2004)

The enrichment of the neon isotopes in a thermoacoustic device is demonstrated. Because the thermal diffusion ratio of neon is small, an apparatus longer than a wavelength was necessary in order to easily observe the separation. The device was modular and extensible, so that arbitrarily large separations could in principle be obtained. The acoustic duct was a series of multiple, identical quarter-wavelength modules with side-branch drivers. In this way, waveforms close to that of a traveling wave were maintained in the duct, despite the high acoustic attenuation caused by the duct's small diameter and large length. The concentrations of the isotopes were measured at one end of the duct using a quadrupole mass spectrometer. For the operating frequency of 227 Hz, the maximum separation gradient obtained was 0.43%/m, and mole fluxes at zero gradient as high as 3 nmol/s were observed. Effects of turbulence, though not observed, are also discussed, and the scaling properties of this method are compared with those of traditional mixture-separation methods. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1687831]

PACS numbers: 43.35.Ud, 43.20.Mv, 43.35.Ty [RR] Pages: 2059–2070

I. INTRODUCTION

Material separation is one of the central challenges of chemical engineering and related industries, and the development of separation techniques remains an active field of research today.[1] Because of the wide range of applications for separation, no single technique addresses all possible needs. Distillation is most widely used because it is, in principle, thermodynamically reversible and therefore offers very high energy efficiency. However, when separating mixtures for which the difference in vapor pressures is small or zero, as in isotope or azeotrope separations, or in cases where distillation would require extremely high or low temperatures or pressures, other techniques can become competitive. Because the capital cost of a distillation device is high for such challenging separations, other methods may be particularly attractive when only small amounts of material are to be processed. The crossover depends on where the cost of a distillation plant becomes comparable to the energy cost of these inefficient methods.

The physical basis for a new separation method based on thermoacoustics was described in several recent articles.[2–5] The thermoacoustic separation process depends on thermal diffusion and viscosity in a binary mixture to produce time-averaged mass fluxes of the mixture's two components in opposite directions. Because both diffusion and viscosity are dissipative, thermoacoustic separation has a low thermodynamic efficiency, on the order of that of conventional thermal diffusion or gaseous diffusion. However, thermoacoustic separation has potential advantages due to its mechanical simplicity and the promise of separation machines quite compact in size.

In this paper, we describe how one can apply the thermoacoustic separation technique in a practical configuration.[6] As a concrete example, we focus on enrichment of the neon isotopes, starting with a sample in which the isotopes are present in their natural abundances. This choice is partly made in deference to history: The first[7] isotope separation by gaseous diffusion was that of neon in 1920, and one of the earliest isotope separations by conventional thermal diffusion was that of neon in 1938 by Clusius and Dickel.[8]

Neon is also a logical choice for this exploration of thermoacoustic isotope enrichment because it is one of the easiest gaseous isotopic systems to separate. The mass difference between the predominant isotopes, ^{20}Ne and ^{22}Ne, is a high fraction of the average mass of the isotopes, resulting in a comparatively large thermal diffusion coefficient. Furthermore, the two predominant isotopes are naturally present[9] in concentrations of about 90.51% (^{20}Ne) and 9.22% (^{22}Ne), more nearly equal than those of other natural binary isotopic mixtures. Thus, no enrichment or preparation of the gas sample was necessary prior to our experiments in order to ensure measurable quantities of both isotopes. Nevertheless, 91–9 ^{20}Ne–^{22}Ne is considerably more challenging[10] to separate than the 50–50 He–Ar mixtures of previous, proof-of-principle experiments,[3,4] because the thermal diffusion ratio is about 50 times smaller.

In the following sections we describe how one can concatenate several modules similar to the experimental device described in Refs. 3 and 4 to separate gas mixtures to a desired degree. A first variation of the method uses a quasi-traveling wave. We outline how one goes about the basic design of such a thermoacoustic separator, we review how one properly quantifies the degree of separation of such a device and each of its modules, and we present an expression relating the number of modules to the separation one can achieve. We then describe an experiment demonstrating neon isotope enrichment by such an extended device, and we compare its measured performance with our calculated expectations. A second variant of the method produces an acoustic impedance higher than that of a traveling wave at equally spaced locations along the duct. Experiments confirm the

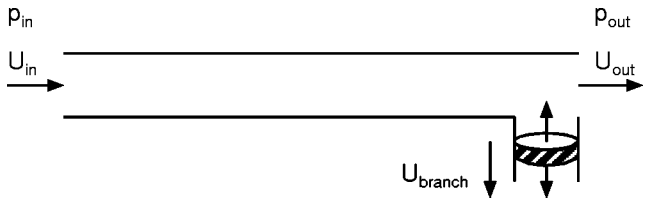


FIG. 1. Assembling a large enough number of modules like this one in series creates a thermoacoustic mixture separator capable of reaching any desired purities.

expected behavior of this variant and demonstrate its modest advantages. The Appendix lists definitions of variables and the results from previous work that are used here.

II. EXTENDING THE SYSTEM

In the experiments[3,4] on He–Ar, the concentrations of the two species could differ by as much as 10% along the tube length of $\sim 1/40$ of a wavelength. For isotopes, the thermal diffusion ratios are generally small so that, to achieve the same degree of separation, either a proportionally longer tube is necessary or one must process the gas in many batches. In the case of laboratory-scale purifications, batch processing of the gas might be reasonable if the number of batch steps is small. However, if the mixture to be processed is so expensive that one cannot simply discard the tailings at the end of each step, then recovery from the tailings requires additional steps and a complex inventory control for the products of each batch. Such batch-mode processing is difficult and tedious.

Hence, thermoacoustic separation in ducts of length greater than a wavelength is advantageous. The method we consider for doing this is to maintain an essentially traveling wave in a duct of arbitrarily great length by providing identical active branches at equally spaced distances along the duct, driven with time phasings appropriate to maintain the amplitude of the wave. Thus, the pressure and volume velocity[11] oscillations at each point between two adjacent drive branches are equal in amplitude but time shifted from the pressure and volume velocity oscillations at the corresponding point between any other pair of adjacent side branches. This is somewhat similar in concept to the ''acoustitron'' of Ref. 12, except that our topology has two ends, comprising a source and a sink for the traveling wave. Were there only one acoustic driver in the system, then the amplitude of the sound wave would decay as the wave traveled along the duct. In our realization of the system, the unit module from which the acoustic network is constructed consists of a length of duct with a tee at one end, one side of the tee leading through a short duct to a sealed bellows, as depicted in Fig. 1. Each bellows is driven lengthwise like a piston by the voice coil of an electrodynamic speaker. At the ends of the concatenation of identical unit modules are similar ducts with similarly driven bellows, one of which starts the wave traveling along the duct while the other actively absorbs it.

The most convenient realization of this scheme has each module one-quarter wavelength long and adjacent speakers phased $-90°$ from one another in time along the duct. This is a practical design because it allows all branches in the acoustic network to be driven with only two electrical signals in quadrature, rather than a larger number of equally spaced signal phasings. This significantly simplifies the electronics needed to drive the array of speakers.

For nonzero spacing between branches, there is always some variation in the amplitudes of pressure and volume velocity as a function of distance along any single module, because of the viscous and thermal attenuations.[5] In addition, the pressure and volume velocity do not have a constant time-phase difference along the duct. As a result, the limiting concentration gradient varies somewhat along the duct. For the oscillating pressure, the boundary condition required on each module is that the amplitude is the same at the beginning and end of the module. The amplitude of volume velocity must also be the same at the beginning and end of the module, but the presence of the branch means that this equality is achieved by adding the volume velocity from the branch to the volume velocity at the end of the duct, with appropriate time phasing. The volume velocity just before the branch is lower than that at the beginning of the duct because of the attenuation along the duct. One result of these two boundary conditions on pressure and volume velocity is that the highest limiting concentration gradient limiting appears near the end of the module just before the branch.

III. DETERMINING THE NUMBER OF MODULES

Before setting out to engineer the acoustics of a thermoacoustic separator, one must assess the kind of performance it will provide. Equations (25) and (36) from Ref. 5, reproduced here in the Appendix, show that the maximum concentration gradient and the maximum efficiency for thermoacoustic separation decrease quickly as the mole fractions of the two components in the binary mixture approach zero and unity, both through their explicit dependence on mole fractions and indirectly through their dependence on the thermal diffusion ratio k_T. The thermal diffusion ratio is approximately proportional to $n(1-n)$, where n is the mole fraction of the component of greatest interest. It is therefore natural to ask whether thermoacoustic separation can ever produce high-purity gases and whether it is at a severe disadvantage in doing so compared to other separation methods. In the case of Eq. (25) in Ref. 5, the factor of k_T implies that the maximum concentration gradient achievable for the desired component falls off approximately[13] as $(1-n)$; this would appear to offer diminishing returns as the enriched gas is further purified. It is more useful, however, to evaluate the separation in terms of the ''separation factor,'' defined in mixture-separation engineering[14] as the ratio of the relative abundances at the beginning and the end of one batch or one stage of processing

$$q=\frac{n_f/(1-n_f)}{n_i/(1-n_i)}, \tag{1}$$

where the subscripts i and f denote initial and final, respectively. For a stepwise separation process, like fractional distillation or gaseous diffusion, q typically turns out to be nearly the same for all steps.[14,15] If several steps are concat-

enated, the total separation factor of the apparatus is equal to the product of the q_j for each step j. For thermoacoustic separation, there are no clearly defined steps, but we may divide the separation duct of total length ℓ into pieces of length Δx. Considering each piece to be a separation step allowed to run to saturation, we can compute the separation factor for the jth piece to be

$$q_j = 1 + \frac{1}{n(1-n)}\left(\frac{dn}{dx}\right)_{\lim} \Delta x. \tag{2}$$

Because most of the n dependence in the expression for the limiting gradient $(dn/dx)_{\lim}$ derives from the factor k_T, q_j is nearly independent of concentration.[13]

In the limit that the duct is subdivided into infinitesimal pieces, we can write for the total separation factor of the duct

$$q_{\text{tot}} = \lim_{\Delta x \to 0} \prod_j q_j = \exp\left[\frac{1}{n(1-n)}\left(\frac{dn}{dx}\right)_{\lim} \ell\right], \tag{3}$$

where n and $(dn/dx)_{\lim}$ are evaluated at any point along the duct. As the duct is lengthened, q_{tot} approaches either 0 or ∞ depending on the sign of $(dn/dx)_{\lim}$, and this corresponds to concentrations n approaching 0 or 1, respectively. Purification of mixed gases thus improves exponentially with the length of the thermoacoustic device, just as it does with the number of stages in distillation or gaseous diffusion plants.

The peak thermodynamic efficiency of thermoacoustic separation is seen from Eq. (36) of Ref. 5 to be approximately proportional to $n(1-n)$, so that it also declines in proportion to the depleted component's concentration. In order to compare methods, it is simplest to consider how much energy must be expended to achieve a given total separation factor q_{tot}. For thermoacoustic separation, the acoustic power needed to obtain a factor q_j is proportional to the length of the duct. As extra lengths of duct are added in order to provide extra factors of q_j, the dissipation will increase linearly with the length. Thermal, viscous, and radiative losses in a Clusius–Dickel column also scale with the length of the column. For a gaseous diffusion plant, mechanical work is lost by free expansion and viscosity at each stage of the plant, so that the free energy dissipated is again proportional to the number of stages with separation factor q_j. Distillation is reversible in principle, but in practice free energy is lost due to pressure drops across the column which are proportional to the number of distillation plates or stages. Cryogenic distillation also suffers from heat leaks which scale as the length of the column. So, even for this ideally reversible process, the losses scale approximately as the number of factors of q_j.

Hence, both for separation factor and for energy efficiency, at high purities the scaling behavior of thermoacoustic mixture separation is the same as that of other mixture-separation systems.

IV. HARDWARE DESIGN CONSIDERATIONS

For this experiment, we developed a $|z| \approx \rho a$ system—a ''nearly traveling-wave system''—to study the separation of neon isotopes starting from a sample of pure neon with natural isotopic abundances. In our notation, z is the complex specific acoustic impedance Ap_1/U_1, ρ is the density of the gas, a is its speed of sound, p_1 is the complex oscillating pressure, U_1 is the complex oscillating volume velocity,[11] and A is the cross-sectional area of the duct. In the experiment, we anticipated needing a change in concentration of each isotope of about 1% across the duct in order that the separation effect be clearly measurable. From Eq. (25) in Ref. 5, reproduced here in the Appendix, it is straightforward to show that

$$\left(\frac{dn}{dx}\right)_{\lim} \approx (\gamma - 1)\frac{2\pi}{\lambda} k_T, \tag{4}$$

for the $|z| \approx \rho a$ system in the boundary-layer limit, where γ is the ratio of the isobaric to the isochoric specific heats and λ is the wavelength. To obtain this expression, we have used the fact that the waves are nearly traveling waves and that F_{trav} and F_{grad} (defined in Ref. 4 and in the Appendix) are of the same order in magnitude, and we have assumed that the amplitude of the sound is high enough that the steady axial diffusion does not appreciably limit the concentration gradient. For monatomic gases like neon, $\gamma = 5/3$. Furthermore, if k_T is small enough that the concentrations of the gas species do not change substantially (i.e., $\Delta n \ll n$) over the span of a single quarter-wavelength module, then k_T itself changes little between the ends of the module. Thus, we can avoid use of Eq. (3) from the previous section and write the separation across a single module as $\Delta n \approx (\gamma - 1)(\pi/2)k_T$. For neon with natural isotopic abundances,[16] $k_T \approx 0.0022$ and we find that our apparatus must be at least four modules long to produce a change of ~ 0.01 in n. We designed our system to contain four quarter-wavelength modules. Because of the additional speakers at either end of the acoustic network to generate or absorb the traveling wave, though, we have a total of six acoustic drivers and approximately five quarter-wavelengths of duct.

For the design of the unit module from which the rest of the device is constructed, one must choose the frequency, the diameter of the duct, and the amplitude of the oscillations in the gas. In view of Eq. (4), one generally wishes to make the frequency as high as possible in order to make the device most compact in size. This has the advantages not only of making it easier to fit the machine in a room or enclosure but also of reducing the inventory of gas inside the machine during operation; the latter is particularly important when the feedstock is either expensive or hazardous. For a fixed total separation factor q_{tot}, though, the number of modules will remain the same no matter what the frequency is. The frequency is limited by either the surface smoothness of the duct or the properties of the drivers used. Any roughness in the inner surface of the duct must be smaller than the viscous and thermal penetration depths, which are themselves proportional to $1/\sqrt{\omega}$. Otherwise, the boundary layer may be destroyed by turbulence and would be difficult to characterize if it existed at all. For high efficiency in producing the sound wave, acoustic drivers must be operated close to their resonance frequency. However, there are constraints to the operating frequencies one can achieve with off-the-shelf components. For example, one may easily lower the resonance frequency of a mass-produced driver by adding mass

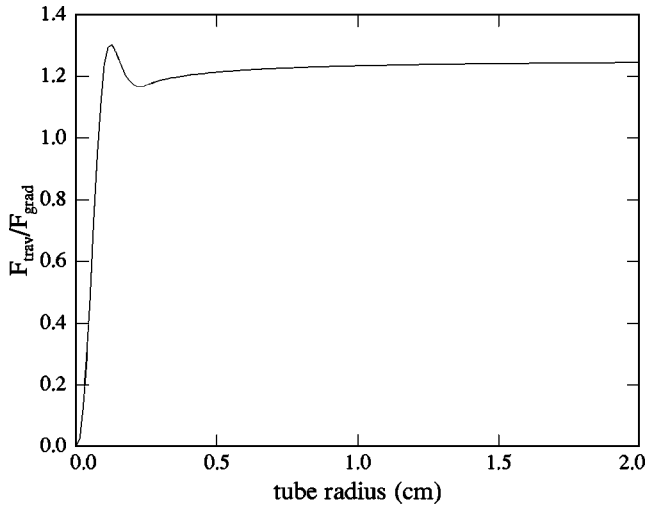


FIG. 2. A measure of the limit that the $|U_1|^2$ remixing imposes on the separation gradient in 80-kPa neon at 227 Hz. When diffusion can be neglected, little advantage is gained by reducing the radius of the duct.

of any type to the fixed moving mass of the actuator. Adding stiffness to raise the resonance frequency, however, is more difficult to achieve in practice: springs and bellows are produced in limited sizes and spring constants, and springs capable of large displacements suffer from surge at high frequencies.

Once the frequency is selected, one must choose a transverse dimension for the duct. We consider here specifically ducts of circular cross section, both because of the availability of such tubing in a variety of sizes and because it is relatively simple to calculate the separation in circular ducts. However, the considerations here will be essentially the same for ducts of arbitrary cross-sectional shape. Because the separation effect depends on the molecules diffusing into and out of the boundary layer, it is necessary for the duct's radius to be greater than the viscous and thermal penetration depths. As was shown by the experiments on He–Ar mixtures,[4] the phasing between p_1 and U_1 which yields the greatest separation is generally not that of a traveling wave, because generally $F_{\text{stand}} \neq 0$. For a given mixture, this optimal phasing also depends on the diameter of the tube. As will be shown in Sec. VIII, for our extended, modular systems the standing-wave component contributes little to the separation even across a single module. Therefore, one might consider reducing the tube diameter to the point at which the optimal phasing is that of a traveling wave. However, the limiting separation per module turns out not to be a maximum for this choice, because both F_{trav} and F_{grad} change with radius in such a way as to *reduce* the limiting concentration gradient as the tube diameter shrinks. Figure 2 demonstrates this for the case of the neon isotopes. The phasing is therefore not an important consideration in choosing the diameter of the duct.

One important factor in choosing the diameter arises instead from steady diffusion. If one can generate high enough velocity amplitudes in the duct that the second-order separation flux is much greater than the mole flux from steady diffusion when the concentration gradient reaches its maximum, then the concentration gradient is limited only by the frequency of operation. Also, it was seen in Ref. 5 that the limiting efficiency of a flow-through separator occurs (in the boundary layer limit) when $|U_1|$ is as high as possible before the onset of conditional turbulence, because the losses due to the remixing by steady diffusion are then insignificant. Thus, to have the highest gradient and best efficiency, one must keep the duct diameter narrow enough that the effect of diffusion is made small for values of $|U_1|$ below the onset of turbulence.

It is generally desirable to have the volume velocity as high as possible without breaking the acoustic drivers and without causing turbulence in a significant fraction of the duct. Also, in order to maintain agreement between the apparatus and its design calculations, one must limit the velocity such that the acoustic approximation may be applied: $\rho a|U_1|/A \approx |p_1| \ll p_m$, where p_m is the mean pressure.

Once the frequency, duct diameter, and maximum velocity are chosen, one can methodically develop the acoustic network. The first step is to design the elementary branched module, solving the equations of acoustics for the required impedance of the branch. For the quasitraveling wave system, a convenient module length is $\lambda/4$. The magnitude of the specific acoustic impedance $|z|$ in the middle of the module will differ slightly from ρa due to the attenuation in a finite-diameter duct; however, this deviation will be small except for tubes with radius of the order of the viscous penetration depth, and it can be corrected through small adjustments to the length of the module. Using the solution for the branch impedance and one's model for the acoustic drivers, one can then solve for the required voltage or displacement of the speaker and its phasing with respect to the pressure (or velocity) at the entrance to the module. The displacements of each branch speaker have the same amplitude, and their phases progress by $-90°$ from branch to neighboring branch along the direction of the traveling wave. From knowledge of the driver properties and of the pressure and velocity exiting the last module, one can design the active termination of the device, which can be considered as absorbing the traveling wave. It is convenient to choose the phasing of this last speaker to be the same as one of the quadrature phasings used in the rest of the series, but the displacement amplitude will generally be different from that of the other speakers. The greatest homogeneity in $|p_1|$ and $|U_1|$ is obtained by choosing the shortest module that makes the speaker phasing the same as that of the last branch speaker. Similarly, one must design the duct to the end speaker that acts as a source for the traveling wave. Again, it is convenient to choose one of the quadrature drive phasings, at the expense of having a third different displacement amplitude. For the greatest homogeneity of the amplitudes of pressure and velocity, one chooses this phase to lead the first branch by 90° and the shortest module for which this is obtained.

Although the separation occurs in the ducts, virtually the entire cost of the apparatus is due to the acoustic drivers. For these we use inexpensive stock parts: Radio Shack 4 1/2-in. RS-1052 woofers[17] and Servometer FC-16 nickel bellows (1-in. outside diameter, 10 convolutions).[18] In order to have a clean, hermetically sealed system with minimal dead volume, the speakers do not contact the neon sample directly.

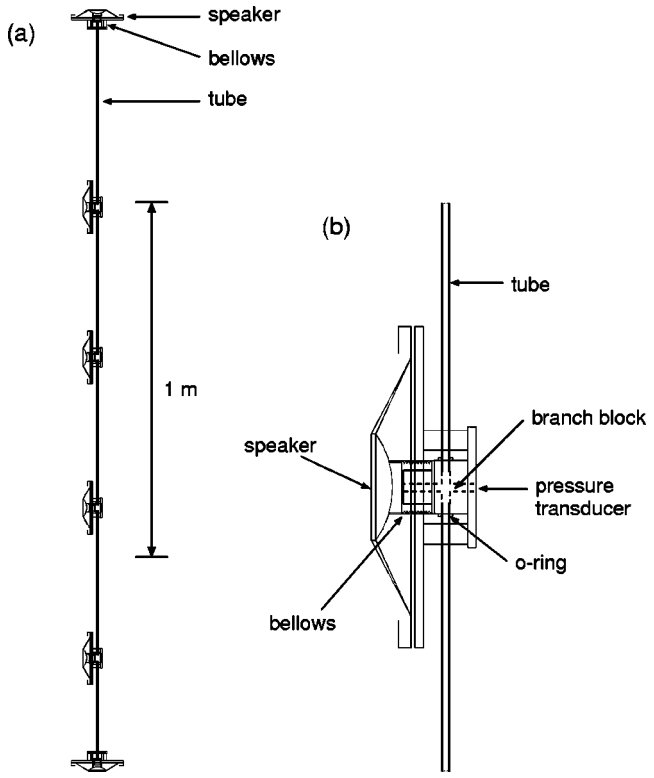


FIG. 3. A schematic of the traveling-wave apparatus. The salient features of the apparatus are shown in (a), while a detailed view of an acoustic driver is shown in (b). (Voice coils and magnet structures not shown.)

Instead, we use the speaker to drive the bellows, sealed on one end and soldered at the other end to a specially machined brass block which serves as either a tee (for side branches) or a straight-through connection (for ends) to the round stainless-steel stock tubing we use for ducts. The brass blocks provide a location for tapping in pressure sensors, thermocouples, and fill or sampling capillaries, and they are rigidly attached to the speakers' frames in order to support the ducts. The brass blocks also protrude into the bellows to eliminate some of their volume, and they present a flat surface as a stop to the sealed end of the bellows in order to prevent the bellows from collapsing beyond its elastic limit when the apparatus is evacuated. To preserve the modularity of the drivers, the tubing is sealed to the brass blocks with O-ring seals. Although the tees could have incorporated 90° bends to create a square coil without creating additional minor loss, we chose to make the system a straight array and mount it on a vertical rail. The linear arrangement made it easy to test different lengths of duct, because then one only needed to slide the speaker mounts along the rail to adjust their spacings. This design is depicted in Fig. 3.

Each bare speaker had a measured resonance frequency of approximately 80 Hz. Besides providing a closed boundary for the gas, the bellows increased the effective spring constant and resonance frequency for the composite driver. The sealed end of the bellows was connected to the speaker's dome by a thin fiberglass tube, which distributed the reaction force of the bellows in a ring far from the center of the dome. This was important, because the dome could easily collapse under a force applied at its center. The tube and the bellows combined to add about 3 g to the bare speaker's own moving mass of 9 g, partly offsetting the increased stiffness. Measurements of the drivers on the apparatus when sealed and filled with the 80-kPa neon used throughout this work showed the new resonance frequency to be about 157 Hz for all the drivers. Little increase in resonance frequency would have been obtained by shortening the fiberglass tube to reduce its mass further.

Although the resonance frequency for our drivers was 157 Hz, we chose to operate them at 227 Hz in the preferred, 90°-spaced mode. For this frequency at room temperature, the wavelength in the neon was 2 m, which for a λ-long apparatus was a convenient vertical clearance to work with in our lab. The duct consisted of stainless-steel tubing with a nominal inner diameter $2R=0.333$ cm, so that $R/\delta_\nu=6$ at 227 Hz. With this system, we were able to generate traveling waves at 227 Hz with $|p_1|$ up to 3350 Pa. This amplitude was sufficient for the concentration gradient to approach within 3% of its limiting value as imposed by the $|U_1|^2$ remixing term in the second-order mole flux.[4] At the highest amplitude, the Reynolds number was 770, based on the tube diameter $2R$ and the velocity $|U_1|/A$.

The quadrature drive signals were generated by two Hewlett-Packard 3325A synthesizers, synchronized by the same external clock in order to keep their relative phase locked. The quadrature signals were fed into a custom-built phase-shift box, consisting of six simple RC networks and isolation transformers, so that the relative phases of the six signals could be adjusted slightly. Outputs of each channel of the phase shifter were then fed to audio amplifiers, allowing us to adjust the voltage amplitude for each of the six speakers individually. The properties of the six acoustic drivers were not perfectly identical and our lumped-element model for the drivers was not perfect, so the ability to adjust both the drive amplitudes and phases was necessary for obtaining nearly ideal traveling waves in our network. The inadequacy of our speaker model had much more impact on the source and sink speakers at the ends of the network than on the side-branch speakers. There, we needed the flexibility to adjust their phases by as much as 20°, depending on the operating frequency. Corrections for the branch drivers, in contrast, were less than 5°. Even without more careful driver construction and analysis, a practical separator would not be strongly affected by such errors in drive voltage, so that the phase-shift network used here would be unnecessary: such an apparatus would consist of many more modules than our prototype does, so that local errors in the acoustics, as at the ends, would decay or be canceled locally and not extend far through the entire device.

For the purpose of creating the desired wave, there were five drive voltage amplitudes and five phases to be adjusted with respect to those of the first speaker, chosen to determine the overall amplitude and phase of the wave. To calculate how close a generated waveform was to our ideal traveling wave, the complex oscillating pressure was measured at each branch.[19] To establish the ideal settings for the complex drive voltages V_1, then, we inverted the 10×10 real matrix of derivative changes in p_1 with V_1 and calculated the matrix product of this inverse with the vector of deviations in p_1

from the ideal waveform. The result was the vector of changes in V_1 necessary to reach the ideal state. This process usually succeeded in producing a very close match to the ideal waveform within two iterations. Because the process of assembling the 10×10 matrix is labor intensive, it is fortunate that it will not be important in the operation of a longer apparatus, as noted above.

V. MEASURING THE NEON ISOTOPE CONCENTRATIONS

While in the He–Ar experiments of Ref. 4 the mole fractions were measured by the speed of sound in two small acoustic resonators, the changes in the speed of sound for the neon experiment would have been near the limit of resolution of that method. For example, a change $\Delta n=0.01$ in the concentration of ^{20}Ne would create a 0.05% change in the speed of sound, 16 times smaller than for a change of $\Delta n=0.01$ in the 50–50 He–Ar system. We therefore decided to measure the neon isotope concentrations directly with a quadrupole mass spectrometer.[20] Because the mass spectrometer requires a high vacuum (HV) system which is disrupted by the transients of opening valves to sample alternately each end of the apparatus, we chose to measure the concentration at one end only, obtaining the end-to-end concentration difference by doubling the difference between measured concentration and natural abundance. The acoustic-absorber end of the device was connected to the HV space of the mass spectrometer through a 5-μm glass capillary[21] cut to 5-cm length. This impedance allowed a mole flux of approximately 2.4×10^{-10} mol/s into the turbopumped HV space, which lowered the pressure in the acoustic network by approximately 1 kPa per day. Since each separation experiment lasted less than 1 day, the change in mean pressure of the apparatus had no noticeable impact on the data.

It was discovered early in our experimentation that the mass spectrometer itself was sensitive to temperature, and most of this sensitivity was associated with the control and detection electronics in spite of its built-in temperature compensation. Therefore, we housed the electronics and part of the mass spectrometer itself inside a sealed styrofoam box, the interior of which was cooled by a chilled water circuit. This thermal isolation held the temperature of the electronics constant to ±10 mK, more than two orders of magnitude better than our lab's climate-control system, making temperature drifts of the mass spectrometer's gains for different masses imperceptible in our data.

We also found it necessary to replace the HV system's original cold-trapped diffusion pump with a dry turbomolecular pump. While each of our experimental curves of concentration versus time took several hours to acquire, the liquid nitrogen level in the original diffusion pump's cold trap would necessarily fall over time, causing a gradual rise in the mass-22 signal until the trap was filled again. This spurious signal was directly correlated with a rise in the CO_2 signal, so it represented a double ionization of the CO_2 (mass 44) rather than a true neon signal. Filling the trap more frequently served only to produce more jumps in the apparent concentration of ^{22}Ne. Thus, the diffusion pump was unacceptable. (We could not use an independent Bayard–Alpert ionization gauge to measure total pressure, because the thoriated-iridium filament produced prodigious quantities of CO_2 which also caused a large, noisy background[22] to the partial pressure measurement of ^{22}Ne.)

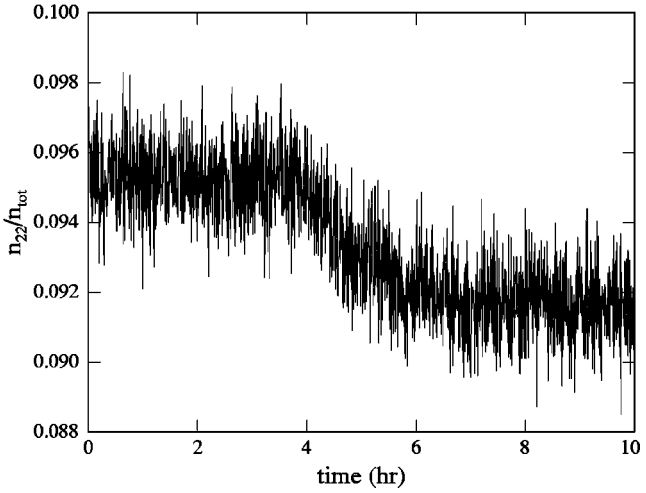


FIG. 4. Raw data from a typical separation experiment. This data set was taken for a traveling wave with $|p_1|=1.1$ kPa and a frequency of 227 Hz. The neon isotopes are at their natural abundances before the sound wave is applied at $t_0\approx4$ h.

Unlike the previous experiments in He–Ar, these experiments allowed continuous measurement of the mole fractions, without stopping the acoustic wave in the separation duct. Figure 4 shows a typical curve for the mole fraction of the ^{22}Ne as a function of time.[23] Although the separation effect is clearly visible, the signal appears noisy. This is partly because of the vertical offset of the graph; the fractional standard deviation is only 1%.

The output of the mass spectrometer is given in terms of the partial pressure for each mass number, and we examined the noise in the recorded partial pressures for several different values of partial pressure and for several different mass numbers. In general we found that the standard deviations were always approximately 1%, regardless of pressure or mass number. This suggests that virtually all of the noise is caused by the electronics and that the shot noise, proportional to $\sqrt{p_i}$, was a negligibly small component of the total noise observed. This is consistent with the fact that this mass spectrometer uses a logarithmic amplifier to span the seven orders of magnitude in partial pressure that it is capable of measuring, rather than by having an adjustable gain amplifier. In the logarithmic amplifier, large input currents are compressed to fit the range of the gain stage, while the gain of the final amplifier is held constant. This results in a constant fractional noise, instead of a fractional noise that is inversely proportional to the input signal as it would be in a detector with variable gain. Although some resolution would be lost even for a low-noise detector of variable gain by the need to digitize the amplifier's analog output signal, the fluctuations in our mass spectrometer were about an order of magnitude greater than the errors from digitization. Thus, much greater signal-to-noise ratio should be possible with other mass spectrometers.

Because of the low signal-to-noise ratio, it was especially important in analyzing the data to constrain the fits as much as possible and thereby reduce their uncertainty. Before each experiment, the mass spectrometer was allowed to run long enough to ensure that it had reached equilibrium, the equilibration time depending on whether the mass spectrometer was just started, was degassed, or had its power cycled momentarily to reboot the control electronics. Baseline partial pressures of masses 20, 22, 40, and 44 were then collected, so that the doubly ionized mass-40 and mass-44 backgrounds could be subtracted from the mass-20 and mass-22 signals, and the starting concentration determined accurately from its average over time. The sound wave was then applied, the start time t_0 was recorded, and the sound was allowed to separate the gases until the mole fraction of ^{22}Ne was not measurably changing. The uncertainty in start times is 20–40 s, determined by the interval between mass spectrometer measurements and by a time lag for the first measurement, originating in the mass spectrometer control software.

In spite of the presence of six reservoirs along the length of the duct (instead of only two as in the He–Ar experiments), the batch-mode limiting can be modeled by an exponential decay to the limiting gradient with a single time constant. Therefore, the data after t_0 were fit to an exponential with only the magnitude and the time constant of the exponential term as free parameters. Because of the quality of the exponential fits in comparison to linear fits near t_0, the time constant was used to determine the initial slope of the separation for comparison with the theory. Often, at the end of an experiment, the valve from the capillary leak to the mass spectrometer's HV chamber was closed, and the residual signals from masses 20, 22, 40, and 44 were remeasured in order to confirm the background pressures of cracked hydrocarbons.

VI. EXPERIMENTAL RESULTS

The first set of experiments was carried out to demonstrate the agreement between the data and theory for this multiple-driver thermoacoustic separator. The difference in steady-state concentration Δn of the ^{22}Ne between one end of the duct and the other was measured as a function of the pressure amplitude $|p_1|$ for a quasitraveling wave at 227 Hz. The results are shown in Fig. 5 and compared with our calculations. These calculations use the measured complex p_1 at each speaker location to calculate the x dependences of p_1, U_1, and $(dn/dx)_{\mathrm{lim}}$ along the duct with the "circular-tube" algorithm of Ref. 4. The concentration difference is then calculated from a numerical integration of $(dn/dx)_{\mathrm{lim}}$ along the duct. Agreement is good, considering the variation in quoted values of the thermal diffusion coefficient $\alpha_T = k_T/n(1-n)$ for neon[16] at room temperature. There is no measurable reduction in the measured separation at the highest $|p_1|$, where turbulent mixing might occur. No turbulent mixing is expected along most of the duct length, because the highest Reynolds number attained in these data is 772, and the duct radius is 6 times δ_ν, putting the motion in the weakly turbulent regime where no significant perturbation of the boundary-layer physics is expected.[24,25] Nevertheless, the result suggests that any turbulence or jetting at the branches, where the acoustic network has tees, does not expand dramatically with the amplitude of sound to adversely affect the separation.

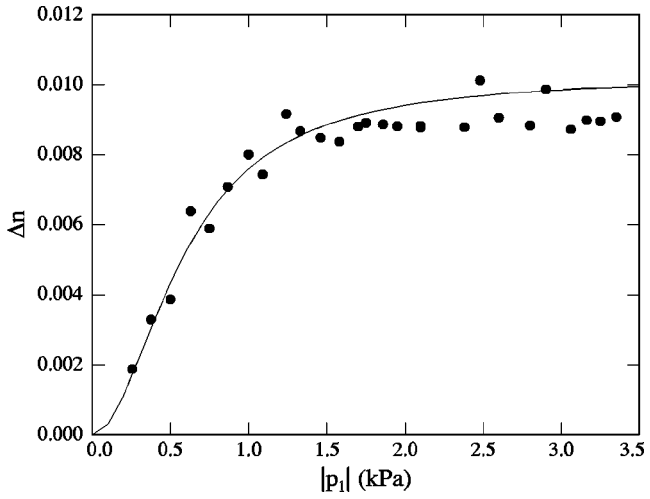


FIG. 5. Total difference in steady-state concentration of ^{22}Ne between ends of the apparatus. Data are for 227 Hz in a duct of total length 2.11 m. Points represent measured values and the curve is the result of a calculation with no adjustable parameters.

Using the time constants from the exponential fits, one can compare the initial mole flux in the apparatus to Eq. (54) of Ref. 4 when $dn/dx=0$. To extract the initial mole flux from a data set, it is tempting to write simply

$$\dot{N}|_0 = \frac{1}{\tau}\frac{\Delta n_{\mathrm{lim}}}{2}\eta V_{\mathrm{res}}, \tag{5}$$

where Δn_{lim} is the total change in concentration across the apparatus in steady-state, τ is the time constant from the exponential fit, η is the molar number density, and V_{res} is the volume of the reservoir at the end of the apparatus where the concentration is measured. However, that identity assumes that all of the volume of the device is in the end reservoirs. For our apparatus, each tubing segment has a volume approximately equal to that of each driver, so most of the volume is in the duct and branch drivers. Therefore, one must add to the volume of the end reservoir the volumes of the duct and branches, weighted by their change in concentration when a finite gradient is established. The end reservoir, like all the acoustic drivers in our system, has a volume of approximately 3.6 cm^3, but the additional, effective volume of the apparatus from the midpoint (where the concentration remains constant) is approximately 7.5 cm^3. The results are plotted against Eq. (54) of Ref. 4 in Fig. 6, where the agreement is seen to be excellent.

VII. SEEKING EFFECTS OF TURBULENCE

Both high mole flux and high concentration gradient demand high acoustic amplitude, but turbulence prevents an arbitrary increase in amplitude. No evidence of turbulence was observed in these measurements, but this section presents a brief exploration of some of the relevant issues and parameter space.

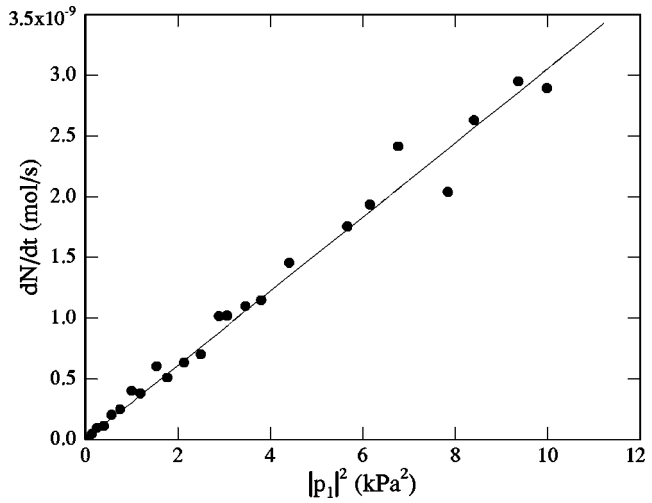


FIG. 6. Theory and data for mole flux of separation at $t=0$, when the components are well mixed, for a frequency of 227 Hz and a duct radius of 1.66 mm.

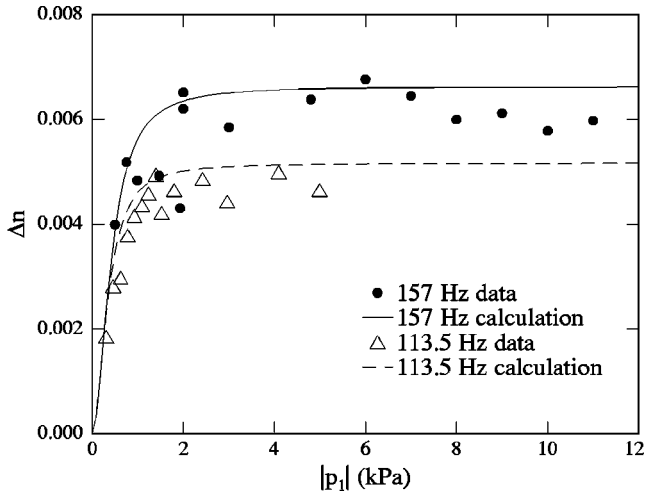


FIG. 7. No evidence of a transition from laminar flow to weak turbulence is seen in the steady-state concentration difference at either 113.5 or 157 Hz.

In the separation duct itself near the necessarily disturbed geometry of each branch, the flow field of the gas is turbulent. This can affect the separation in two ways. First, the turbulence causes a pressure drop near the branch—the ''minor loss''—so that the branch speaker will have to be driven harder than calculations assuming laminar flow would predict. Second, the turbulence destroys the boundary layer and remixes the gas constituents in the vicinity of the branch. Since this remixing takes place very close to the location where the separative effect should be strongest, and the turbulently disturbed region probably grows with increasing amplitude, the true separation across the duct could be significantly lower than the laminar theory predicts.

Turbulence may hinder operation of the nearly traveling-wave configuration in a more spatially extended way as well. If, for a fixed operating frequency of the apparatus, the duct is narrower in diameter than approximately 5 times the viscous penetration depth δ_ν, then for low-amplitude operation the oscillating motion of the gas will be laminar.[24,26,27] At velocity amplitudes high enough to make the Reynolds number 1000–2000, though, the oscillating flow in the duct can cross into the conditionally turbulent regime, destroying the boundary layer wherever such high velocities occur. This will eliminate the separative effect from the beginning of each module out to the point along the duct at which the local velocities are low enough for the gas motion to be laminar. For traveling-wave systems in which the module length is less than $\lambda/2$, the velocity amplitude is relatively constant so that the entire module will become turbulent at a pressure amplitude only slightly larger than that at which turbulence first appears. This crossover to conditional turbulence should also occur when the duct's radius is larger than $\sim 5\delta_\nu$, but such systems will be in the weakly turbulent regime at low amplitudes and will cross over to conditional turbulence only at significantly higher amplitudes of velocity compared to the laminar case.

To experimentally explore the possibility of turbulence destroying thermoacoustic separation, we studied the performance of the apparatus at two other frequencies, 113.5 and 157 Hz. For these frequencies, we chose to abandon the preferred $-90°$ phase-shift condition between adjacent branches, instead using our phase-shift networks to create the appropriate phasing that produced traveling waves. This was done in order to avoid changing the duct lengths. At these lower frequencies, Reynolds numbers up to 2500 were reached, but operating conditions were always close to the region of crossover from laminar to weakly turbulent flow.[24] At the highest amplitudes $|p_1|/p_m > 1/10$, so that the validity of the acoustic approximation used in calculating the separation is questionable. Nevertheless, all these lower-frequency data (Fig. 7) show no abrupt decline in the separation at high amplitudes. Although these data do not rule out effects of turbulence, they suggest that the penalty in separation for having weakly turbulent flow is small. Also, the duct sizes for which a transition to conditional turbulence would be easiest to achieve are so small, of order $R/\delta_\nu < 3$, that the separation would be disadvantaged by the effect shown in Fig. 2.

VIII. THE HIGH-IMPEDANCE APPROACH

In order to mitigate the sources of turbulent mixing, it may be useful to consider configurations of concatenated branched modules with a specific acoustic impedance higher than ρa, at least near the middle of each module. If the acoustic impedance is increased, the volume velocity—and, therefore, the Reynolds number—will be lower for the same value of acoustic power. From Eq. (54) in Ref. 4, one also sees that increasing $|Z| = |p_1|/|U_1|$ as the product $|p_1||U_1|$ is held fixed can lead to larger values of the concentration gradient because it decreases the relative importance of the $|U_1|^2$ remixing process compared to the mole flux of separation. Thus, increasing $|Z|$ seems like a promising way to improve the performance of thermoacoustic separation. Although this is true, one does not have as much flexibility in choosing $Z(x)$ as one might expect.

Ideally, one would like to adjust $|Z|$ to be as high as possible, and to choose the phase $\theta = \arg(Z)$ to be that which would produce the largest concentration gradient, i.e., $\theta = \tan^{-1}(F_{\text{stand}}/F_{\text{trav}})$. To do this, there are two parameters of

each module one may adjust: the module's length L and the temporal phase lag $\Delta\phi$ between p_1 at the beginning and end of the module. (Note that the volume velocity must share this phase lag; otherwise, the phase θ between p_1 and U_1 would change incrementally at the beginning of each successive module, eventually reversing the direction of the mole fluxes of the components and defeating the purpose of the concatenation of modules.)

In order to relate the magnitude and phase of the acoustic impedance in the middle of the duct to L and $\Delta\phi$, we need the solution to the coupled first-order equations for $p_1(x)$ and $U_1(x)$ in the duct.[28] Applying the boundary condition for the duct segment $p_{\text{out}}=p_{\text{in}}e^{\imath\Delta\phi}$ in the equation for $p_1(x)$, one finds that

$$p_{\text{in}}=\frac{-\imath Z_0 U_{\text{in}}\sin kL}{e^{\imath\Delta\phi}-\cos kL}, \tag{6}$$

$$p_1(x)=-\imath Z_0 U_{\text{in}}\left[\frac{\sin kL\cos kx}{e^{\imath\Delta\phi}-\cos kL}+\sin kx\right], \tag{7}$$

$$U_1(x)=U_{\text{in}}\left[\cos kx-\frac{\sin kL\sin kx}{e^{\imath\Delta\phi}-\cos kL}\right], \tag{9}$$

where k is the complex wave number, $Z_0=\omega\rho_m/kA(1-f_\nu)$ is the acoustic impedance of a traveling wave in the duct of area A, and f_ν is the complex correction for viscosity. In the middle of the duct we set $x=L/2$ to find

$$z(x=L/2)\equiv z_{\text{mid}}=-z_0\tan(kL/2)\cot(\Delta\phi/2), \tag{10}$$

where the $z_i=Ap_1/U_1$ are specific acoustic impedances. Because both z_0 and k are generally complex, z_{mid} typically has a complex value. From this expression, we see that changing the phase shift $\Delta\phi$ across the duct will alter the magnitude of the impedance in the middle of the duct but not the phase. Changing the length of the duct, however, alters both the magnitude and the phase of the impedance in the middle. Nevertheless, it is not possible to arbitrarily choose both $|z|$ and θ: for example, no combination of L and $\Delta\phi$ allows one to obtain $\theta\geqslant 0$ in the middle of the tube. For ducts wide enough to be well represented by calculations in the boundary-layer limit, $F_{\text{stand}}/F_{\text{trav}}>0$ so that it is not possible to make the phasing at the center of the tube optimal.[2,4] This is not an important restriction in choosing the radius of the duct, because both analytical and numerical studies show that the standing-wave component never contributes significantly to the separation across a module.

As for the nearly traveling-wave system described above, $|p_1|$, $|U_1|$, θ, and the limiting concentration gradient will vary along each module of a configuration designed to provide $|z|>\rho a$. Alterations of L and $\Delta\phi$ to generate these higher-z configurations effectively do so by adding in a traveling-wave component opposed to the direction of the wave in the $z\approx\rho a$ system. The velocity of this added wave attenuates as it travels in the opposite direction from each branch, ensuring that the velocity has its minimum somewhere away from the ends of the module. This tends to push the maximum of $|z|$ and of the concentration gradient away from the ends of the module as desired.

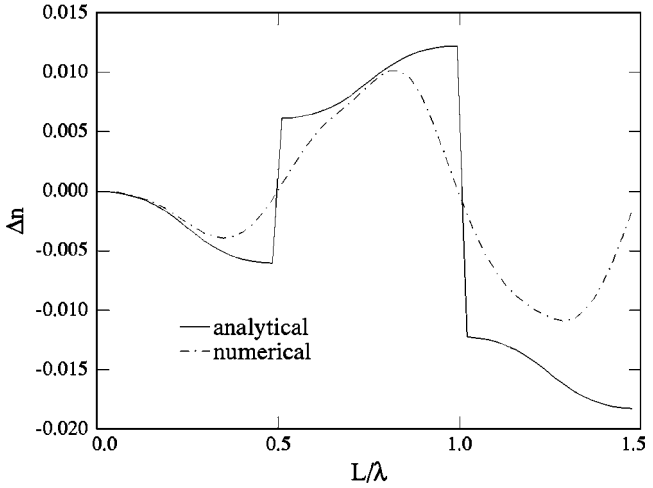


FIG. 8. The separation Δn per module, as a function of module length L, calculated for the inviscid analytical model (solid curve) and for the full numerical model (dash-dot curve) in the boundary-layer limit. The phasing across the branch is fixed at $\Delta\phi=-90°$.

There is a large parameter space to analyze in determining what is the best higher-z configuration to use in a separation device. One must first decide whether it is most important that the device be as short as possible, and thus to generate the highest gradient dn/dx, or whether it is more important to maximize the separation Δn per module of the acoustic network. Here, we make the latter choice, because the number of active branches is probably most important in terms of a separator's capital cost. One may then look for the combination of L and $\Delta\phi$ that provides the highest Δn.

We have studied this problem both numerically and with a simplified analytical calculation.[29] Plots of both numerical and analytical calculations of Δn versus L with $\Delta\phi$ held fixed at $-90°$ are shown in Fig. 8. For module length $L>\lambda/4$ the curves begin to diverge (although they remain similar qualitatively), with the analytical calculation predicting higher Δn because it neglects viscosity and thermal conductivity in the wave number. For $L>\lambda/2$, dn/dx and $\dot{N}_2$ vary spatially along the tube in a complicated way, possibly yielding locations where there is a bottleneck in the separation flux. If the parameter space is therefore limited to $L\leqslant\lambda/2$, one finds that Δn will be maximized as $L\to\lambda/2$ and $\phi\to 0$. The resulting Δn in that limit is, for the analytical calculation, a factor of 2 higher than for the preferred, traveling-wave system we considered in the previous sections.

The actual improvement in Δn that can be achieved will be smaller than this, because it is dangerous to design L too close to $\lambda/2$. For L near $\lambda/2$, a small decrease in the speed of sound (arising from, say, a change in temperature or gas composition) would cause L to become greater than the intended $\lambda/2$. This would lead to a change in both magnitude and sign for the separation in that module, possibly defeating the purpose of the device.

In order to test our understanding of the high-z approach, we measured the separation as a function of $|p_1|$ for $|z_{\text{mid}}|\approx 2.5\,\rho a$ at 227 Hz. This configuration was achieved by keeping all the duct lengths the same but phasing adjacent

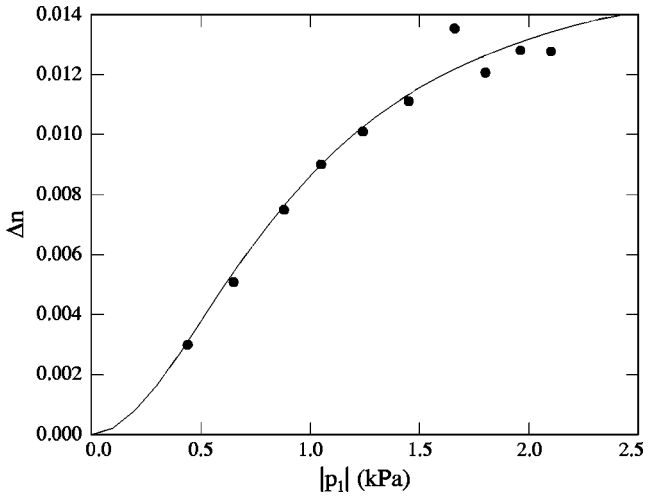


FIG. 9. The difference in steady-state concentration of ^{22}Ne across the apparatus for $|z_{\mathrm{mid}}| \approx 2.5\ \rho a$ in the middle of each module.

branch speakers $-45°$ apart in time, instead of the $-90°$ spacing of the nearly traveling-wave system. The data from this study are shown in Fig. 9, where agreement with calculations is again seen to be quite good. As $|p_1|$ is increased, calculations show that Δn should approach a limiting value of ~ 0.015, or approximately 50% higher than for the nearly traveling-wave configuration.

IX. CONCLUSIONS

This paper shows that thermoacoustic mixture separation, discovered only in the past decade, is extensible to arbitrarily high purities by maintaining a quasitraveling wave through an arbitrarily large number of wavelengths of sound, using a periodic, modular series assembly of tubes and drivers. In the proof-of-principle ^{22}Ne enrichment described here, an enrichment from the natural abundance of 0.092 to 0.097 was achieved at one end of the apparatus with 5 quarter-wavelength modules using $|z| \approx \rho a$. From the point where the mixture consisted of its natural abundances, this enrichment required approximately 2.5 modules; to obtain 0.90 ^{22}Ne from a reservoir of natural abundance would require only 148 modules, and to obtain 0.99 ^{22}Ne would require only 232 modules.

A variant of the method, using a spatially periodic wave with regions of enhanced acoustic impedance, promises modest, factor-of-2 reduction in the number of modules required for a given purity. Other interesting, quantitative physics issues that could lead to modest changes in expected performance include the transition to turbulence, which was not encountered in the present measurements, and the remixing effect of streaming.[4]

Further analysis and experimentation are also required to develop practical steady-flow separators in which a nonzero time-independent mole flux and nonzero concentration gradient are present simultaneously, and to learn whether single devices can achieve high throughput by use of stack-like parallel arrays of channels without counterflowing streaming in different channels causing unacceptable remixing.

The modular apparatus can be coiled, with either curved or straight tubular modules, to form a compact system, one in which each same-phase set of periodically arrayed acoustic drivers can share common electromechanical transduction hardware if the length of each turn of the coil equals one wavelength of sound in the gas. However, to fully enjoy the low-cost, high-reliability promise of thermoacoustics, the thousands of joints that would be present in a straightforward multiplication of our 5 modules should be eliminated. Perhaps future work will demonstrate a good way to drive a peristaltic wave in the cross-sectional area of a long, coiled, flexible-walled tube without joints.

ACKNOWLEDGMENTS

The authors are grateful to Mark Hollander for his assistance in improving our vacuum system. We also thank Scott Backhaus for inducing us to consider the effect of standing-wave components. This work was supported by the Office of Basic Energy Sciences in the U.S. Department of Energy's Office of Science.

APPENDIX: RESULTS FROM EARLIER WORK

Useful results and definitions from Refs. 2, 4, and 5 are collected here, expressed in terms of mole fractions and mole fluxes. For a discussion in terms of mass fractions and mass fluxes, see those references.

The separation flux to second order is given by combining Eqs. (41), (44), and (56) of Ref. 4

$$\dot{N}_H = \frac{\delta_\kappa}{4r_h}\frac{\gamma-1}{\gamma}\frac{k_T}{R_{\mathrm{univ}}T_m}|p_1||U_1|[F_{\mathrm{trav}}\cos\theta + F_{\mathrm{stand}}\sin\theta] + \frac{\delta_\kappa}{4r_h}\frac{\rho_m|U_1|^2}{\omega A M_{\mathrm{avg}}}F_{\mathrm{grad}}\frac{dn_H}{dx} - NAD_{12}\frac{dn_H}{dx}. \tag{A1}$$

In the boundary-layer approximation, the Fs are given by

$$F_{\mathrm{trav}} = \frac{\sigma\sqrt{\sigma L} - \sqrt{\sigma} - \sigma\sqrt{L}(\delta_{\kappa D}/\delta_\kappa + \delta_{D\kappa}/\delta_\kappa)}{(1+\sqrt{L})[(1+\sigma)(1+\sigma L)+\varepsilon\sigma]}, \tag{A2}$$

$$F_{\mathrm{stand}} = \frac{-\sigma\sqrt{\sigma L} + \sqrt{\sigma} - \sigma\sqrt{L}(\delta_{\kappa D}/\delta_\kappa + \delta_{D\kappa}/\delta_\kappa)}{(1+\sqrt{L})[(1+\sigma)(1+\sigma L)+\varepsilon\sigma]}, \tag{A3}$$

$$F_{\text{grad}}=\frac{\sqrt{\sigma}L(1-\sigma^2)(1+\sqrt{L})+\varepsilon\sqrt{\sigma}(\sqrt{L}-1)+[(\sigma^2-1)L+\varepsilon\sigma\sqrt{L}](\delta_{\kappa D}/\delta_\kappa+\delta_{D\kappa}/\delta_\kappa)}{(1+\sqrt{L})[(1+\sigma)(1+\sigma L)+\varepsilon\sigma][(1-\sigma)(1-\sigma L)-\varepsilon\sigma]/\sigma} \tag{A4}$$

with

$$\delta_{\kappa D}^2=\tfrac{1}{2}\delta_\kappa^2[1+(1+\varepsilon)/L+\sqrt{[1+(1+\varepsilon)/L]^2-4/L}], \tag{A5}$$

$$\delta_{D\kappa}^2=\tfrac{1}{2}\delta_\kappa^2[1+(1+\varepsilon)/L-\sqrt{[1+(1+\varepsilon)/L]^2-4/L}], \tag{A6}$$

$$\varepsilon=\frac{\gamma-1}{\gamma}\frac{k_T^2}{n_H(1-n_H)}. \tag{A7}$$

The more complicated Bessel-function forms of the Fs, accurate for circular tubes of arbitrary radius, are described (but not displayed) in Ref. 4.

Equation (A1) describes the mole flux for any value of the concentration gradient dn_H/dx. Our experiments typically start with $dn_H/dx=0$ and proceed toward the limiting state where $\dot{N}_H=0$. In this limiting state, the concentration gradient is given by Eq. (25) of Ref. 5

$$\left(\frac{dn_H}{dx}\right)_{\text{lim}}=\frac{(F_{\text{trav}}\cos\theta+F_{\text{stand}}\sin\theta)(\gamma-1)k_T|p_1||U_1|/A\gamma p_m}{4D_{12}r_h/\delta_\kappa-F_{\text{grad}}|U_1|^2/A^2\omega}. \tag{A8}$$

TABLE I. Variables used in the Appendix. Units are MKS.

A	area of duct (m^2)
c_p	isobaric heat capacity per unit mass (J/kg·K)
D_{12}	binary mass diffusion coefficient (m^2/s)
$\dot{E}$	acoustic power (W)
k	thermal conductivity (W/m·K)
k_T	thermal diffusion ratio
L	$k/\rho_m c_p D_{12}$
M_{avg}	mixture average molar mass (kg/mol)
$\dot{N}_H$	mole flux of heavy component (mol/s)
N	mixture molar density (mol/m^3)
n_H	mole fraction of heavy component
p_m	mean pressure (Pa)
$\|p_1\|$	amplitude of oscillating pressure (Pa)
R_{univ}	universal gas constant (J/mol·K)
r_h	hydraulic radius of duct (m), A/Π
T_m	mean temperature (K)
$\|U_1\|$	amplitude of oscillating volume velocity (m^3/s)
x	coordinate along sound-propagation direction (m)
γ	ratio of isobaric to isochoric specific heats
δ_ν	viscous penetration depth (m), $\sqrt{2\mu/\omega\rho_m}$
δ_κ	thermal penetration depth (m), $\sqrt{2k/\omega\rho_m c_p}$
δ_D	mass-diffusion penetration depth (m), $\sqrt{2D_{12}/\omega}$
θ	phase angle by which p_1 leads U_1
μ	dynamic viscosity (kg/m·s)
Π	duct perimeter (m)
ρ_m	mean density (kg/m^3)
σ	Prandtl number, $\mu c_p/k$
ω	angular frequency (s^{-1})

Acoustic power per unit length consumed by the mixture-separation process is given by Eqs. (19)–(22) in Ref. 5, but an excellent approximation is that given for a pure gas by Eq. (5.10) in Ref. 25. In the boundary-layer approximation, the latter becomes

$$\frac{d\dot{E}_2}{dx}=-\frac{\Pi\delta_\nu}{4}\rho_m\left|\frac{U_1}{A}\right|^2\omega-\frac{\Pi\delta_\kappa}{4}\frac{|p_1|^2}{\gamma p_m}(\gamma-1)\omega. \tag{A9}$$

The thermodynamic efficiency of thermoacoustic mixture separation is the ratio of the rate at which the Gibbs free energy of the mixture increases to the rate of acoustic power consumed. Reference 5 shows that this efficiency is highest when dn_H/dx is half of the limiting value given in Eq. (A8). Using this value of concentration gradient and several other simplifying assumptions yields the upper bound on the efficiency given by Eq. (36) in Ref. 5

$$\eta_{\text{best}}\approx\frac{\varepsilon}{4}\frac{2(1+\sqrt{\sigma L})}{\sqrt{L}(1+\sqrt{L})^2(1+\sigma)}. \tag{A10}$$

Other variables are defined in Table I.

[1] *Encyclopedia of Separation Technology*, edited by D. M. Ruthven (Wiley, New York, 1997).

[2] G. W. Swift and P. S. Spoor, "Thermal diffusion and mixture separation in the acoustic boundary layer," J. Acoust. Soc. Am. **106**, 1794–1800 (1999); **107**, 2229(E) (2000); **109**, 1261(E) (2001).

[3] P. S. Spoor and G. W. Swift, "Thermoacoustic separation of a He–Ar mixture," Phys. Rev. Lett. **85**, 1646–1649 (2000).

[4] D. A. Geller and G. W. Swift, "Saturation of thermoacoustic mixture separation," J. Acoust. Soc. Am. **111**, 1675–1684 (2002).

[5] D. A. Geller and G. W. Swift, "Thermodynamic efficiency of thermoacoustic mixture separation," J. Acoust. Soc. Am. **112**, 504–510 (2002).

[6] D. Geller, G. Swift, and S. Backhaus, "Method and apparatus for separating mixtures of gases using an acoustic wave," U.S. Serial No. 10/238,250, filed with the U.S. Patent Office on 10 September 2002; allowed 13 January 2004.

[7] D. Massignon in *Topics in Applied Physics: Uranium Enrichment*, Vol. 35, edited by S. Villani (Springer, New York, 1979), pp. 55–56.

[8] K. Clusius and G. Dickel, "Neues Verfahren zur Gasentmischung und Isotopentrennung," Naturwissenschaften **26**, 546(L) (1938).

[9] See *Handbook of Chemistry and Physics, 64th ed.* (CRC Press, Boca Raton, FL. 1984), p. B-235. Besides the two main isotopes, ^{21}Ne is present at a concentration of 0.27%.

[10] As a measure of the difficulty in separating the neon isotopes compared to He–Ar mixtures, the 157-Hz neon data described below can be compared with data we obtained at the same frequency after filling this apparatus with a sample mixture of 55–45 He–Ar, which has the same sound speed as neon. In He–Ar, with a pressure amplitude of 2650 Pa, the apparatus produced concentration differences $\Delta n\sim 0.4$ across the duct—more than 50 times larger than the Δn obtained with 91–9 ^{20}Ne–^{22}Ne mixtures at the same operating frequency.

[11] Acousticians use the term "volume velocity" for what engineers call "volume flow rate."

[12] C. C. Lawrenson, L. Dwynn Lafleur, and F. Douglas Shields, "The solution for the propagation of sound in a toroidal waveguide with driven walls (the acoustitron)," J. Acoust. Soc. Am. **103**, 1253–1260 (1998).

[13] The thermal diffusion ratio is in general nearly proportional to $n(1-n)$. He–Ar mixtures, for example, have $k_T=0.38n^{0.8}(1-n)^{1.2}$, based on fits to the data of Atkins *et al.* [Proc. R. Soc. London, Ser. A **172**, 142–158

(1939)]. For this reason, the literature often defines another parameter $\alpha_T \equiv k_T/n(1-n)$, called the thermal diffusion coefficient. The thermal diffusion coefficient is a function of temperature but has only a weak dependence on the concentrations.
[14] Stelio Villani, *Isotope Separation* (American Nuclear Society, LaGrange Park, IL, 1976).
[15] H. London, *Separation of Isotopes* (Newnes, London, 1961).
[16] S. Weissman, "Self diffusion coefficient of neon," Phys. Fluids **16**, 1425–1428 (1973).
[17] RadioShack Corp., 300 West Third Street, Suite 1400, Fort Worth, TX 76102. http://www.radioshack.com
[18] Servometer, 501 Little Falls Road, Cedar Grove, NJ 07009; http://www.servometer.com
[19] 8510B pressure transducer, Endevco, San Juan Capistrano, CA; http://www.endevco.com; SR830 lock-in amplifier, Stanford Research Systems, Sunnyvale, CA; http://www.srsys.com
[20] RGA100, Sanford Research Systems, Sunnyvale, CA, http://www.srsys.com
[21] Polymicro Technologies, LLC, 18019 N. 25th Ave., Phoenix, AZ 85023; http://www.polymicro.com
[22] J. F. O'Hanlon, *A User's Guide to Vacuum Technology*, 2nd ed. (Wiley, New York, 1989), p. 141.
[23] The reader may notice that the equilibrium concentration in Fig. 4 before separation appears to be 9.53% rather than the 9.22% quoted in Ref. 9. We attribute this to an error in the calibration of the mass spectrometer.
[24] M. Ohmi, M. Iguchi, K. Kakehashi, and T. Masuda, "Transition to turbulence and velocity distribution in an oscillating pipe flow," Bull. JSME **25**, 356–371 (1982).
[25] G. W. Swift, *Thermoacoustics: A Unifying Perspective for some Engines and Refrigerators* (Acoustical Society of America, Melville, NY, 2002).
[26] U. H. Kurzweg, E. R. Lindgren, and B. Lothrop, "Onset of turbulence in oscillating flow at low Womersley number," Phys. Fluids A **1**, 1972–1975 (1989).
[27] M. Hino, M. Sawamoto, and S. Takasu, "Experiments on transition to turbulence in an oscillatory pipe flow," J. Fluid Mech. **75**, 193–207 (1976).
[28] Reference 25, Eqs. (4.54) and (4.70) with $dT_m/dx = 0$.
[29] The analytical calculation of Δn across a module of arbitrary L and $\Delta\phi$ becomes tractable when turbulence and steady diffusion are ignored and the wave number is taken to be real.

Measurements of the resistance of parallel-plate heat exchangers to oscillating flow at high amplitudes

Ray Scott Wakeland and Robert M. Keolian[a)]
The Pennsylvania State University, Graduate Program in Acoustics, P.O. Box 30, State College, Pennsylvania 16804-0030

(Received 29 July 2003; revised 30 January 2004; accepted 31 January 2004)

Measurements of the acoustic resistance of parallel-plate heat exchangers are reported. The resistance is measured for two identical exchangers separated by a small gap, and also by a large gap. High amplitude deviations from linear theory are analyzed in terms of a minor loss coefficient. Results are compared to theoretical predictions made in a previous article. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1701901]

PACS numbers: 43.35.Ud, 43.25.Ed, 47.60.+I [RR] Pages: 2071–2074

I. INTRODUCTION

The authors have recently completed a study of heat transfer between identical parallelplate heat exchangers in oscillating flow.[1] Pressure data, collected during the same experimental runs in which the heat transfer data were collected, are reported here, with emphasis on the nonlinear, high-amplitude resistance to oscillating flow. In a previous article,[2] we developed equations for calculating effective exit-flow minor loss coefficients for nonuniform, time-varying, oscillating flows, and applied the results to velocity profiles found between parallel plates in thermoacoustic devices. The present measurements provide some experimental support for that previous theoretical work.

II. EXPERIMENTAL APPARATUS

The experimental apparatus is described in detail in Refs. 1 and 3. Essentially, a pair of identical parallel-plate heat exchangers is inserted into a high-amplitude, low-frequency, oscillating flow in air at atmospheric pressure, and the pressure drop across the exchangers is measured. Reference 3 contains validation of the pressure measurement system used for the present measurements.

The heat exchangers used in these measurements are made from flat extruded aluminum tubing of a type used in automotive air-conditioner condensers. The tubes are 2.0 mm $\times$22.0 mm in external cross section, having rounded ends with a nominal radius of 1.0 mm. Each 292 mm$\times$292 mm exchanger is made up of 35 of these tubes separated by open spaces of 6.35 mm, for a center-to-center tube spacing of 8.35 mm and hydraulic radius of $r_h = y_0 = 3.175$ mm. (The hydraulic radius for parallel plates is half of the plate separation. This half spacing is often referred to as y_0 in thermoacoustics, following the notation of Swift.[4] Figure 1 shows a few of the tubes in cross sections, with dimensions labeled.) The total wetted perimeter of the tubes is $\Pi = 20.45$ m, and the porosity σ (exchanger void volume divided by total exchanger volume) is 0.76. A piece of foam is used to insulate the exchanger manifolds from each other, and to establish the separation $2x_{\mathrm{hx}}$ of adjacent ends of the exchanger tubes. Results are reported for two sizes of spacer: ''small gap,'' $2x_{\mathrm{hx}} = 5.5$ mm; and ''large gap,'' $2x_{\mathrm{hx}} = 54.4$ mm. The exchangers are aligned so that, if one could look axially down the duct, the tubes of the closer exchanger would lie directly in front of the tubes of the far exchanger, with a maximum amount of ''free flow area.'' Quantities used in the analysis include frequency f, angular frequency $\omega = 2\pi f$, peak velocity amplitude u_1 of the gas in the duct, the higher amplitude $u_{\mathrm{hx}} = u_1/\sigma$ at locations within the exchangers, the kinematic viscosity ν of the air, viscous penetration depth $\delta_\nu = \sqrt{2\nu/\omega}$, total frontal area A_{fr} of the exchangers (which is also the cross-sectional area of the duct), and volume velocity $U_1 = A_{\mathrm{fr}} u_1$. A mercury barometer is used to determine the density ρ_m of the air, with $\rho_m = 1.145$ kg/m^3 for these experiments.

Because the pressure drop across the exchangers is very small, the useful range of frequencies is much more limited for the pressure measurements than for the heat transfer measurements. A pair of bellows are used as part of the apparatus that drives the air back and forth between the exchangers, as described in detail in Ref. 3. In the direction of high frequency, as the first resonance of the bellows is approached, extraneous bellows motions affect the phase of the gas displacement, so that it can no longer be accurately inferred from the end-plate motion. Because the inertial component of the pressure increases with frequency squared (for constant displacement amplitude), the phase shift due to the resistive component becomes very small, so that even a small error in phase produces a large error in the calculated resistance. In the low frequency direction, at some point the total pressure simply becomes too small to measure for the parallel plate exchangers, which present very little resistance to flow. The reliable range of frequencies is 1.5–6 Hz, corresponding to $3.45 > y_0/\delta_\nu > 1.72$. A complete description of the measurements used to determine reliability is included in Wakeland's dissertation.[5]

The oscillating flows within the exchangers are very likely laminar. The Reynolds numbers shown on the upper axes of Figs. 2–5 are *peak* Reynolds numbers, $\mathrm{Re}_{\mathrm{hx}} = u_{\mathrm{hx}} 4y_0/\nu$, in terms of hydraulic diameter ($=4y_0$), with a maximum value of 2020. The acoustic, oscillating-flow Reynolds number, $\mathrm{Re}_{\mathrm{ac}} = u_{\mathrm{hx}}\delta_\nu/\nu$, depends on frequency. Its

a)Electronic mail: keolian@psu.edu

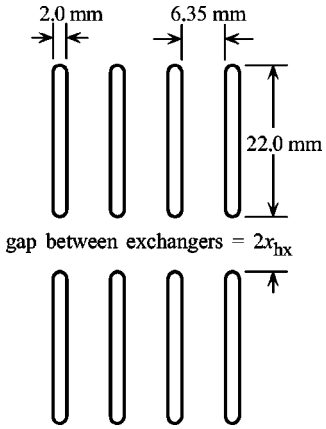


FIG. 1. Dimensions of the heat exchangers.

maximum value for any of these data points is 160. The flows within the exchanger, then, would be expected to be strictly laminar if the exchangers were long (Ref. 4, Sec. 7.2). The shortness of the exchangers puts this conclusion in some doubt, but the low value of $\mathrm{Re}_{hx}<2020$ itself also suggests laminar conditions. The peak Reynolds number within the duct is 35 000, but the ratio $\sqrt{A_{fr}}/\delta_\nu>26$ puts these flows in the ''conditionally turbulent'' regime, so the boundary-layer calculation of resistance in the duct is appropriate.

III. RESULTS AND ANALYSIS

Small-gap results are shown in Figs. 2 and 3, with large-gap results in Figs. 4 and 5. Data are presented in terms of acoustic resistance ($R=\Delta p/U_1$) versus volume-velocity amplitude U_1, where Δp is the *resistive* component of the pressure difference across the test section, i.e., the component of the total pressure drop that is in phase with velocity. The total, measured resistance (divided by two) is shown by the open circles for various values of y_0/δ_ν.

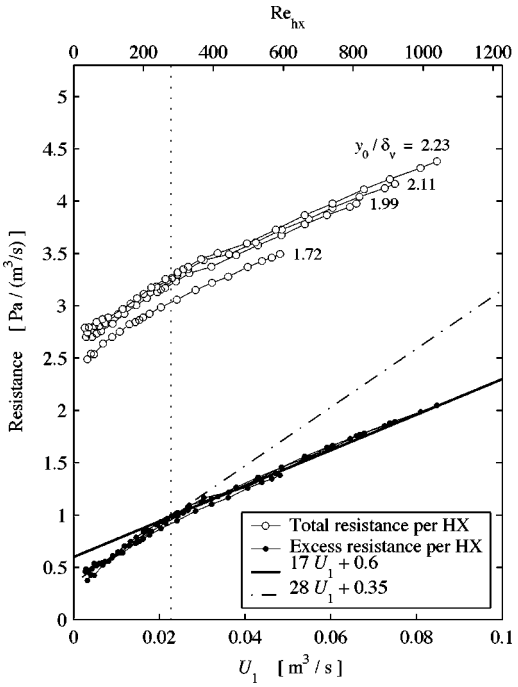


FIG. 3. Measured acoustic resistance per exchanger at 1.5, 2, 2.25, and 2.5 Hz for small exchanger separation ($2x_{hx}=5.5$ mm). There seems to be a knee in the data at around $U_1=0.023\ \mathrm{m}^3/\mathrm{s}$.

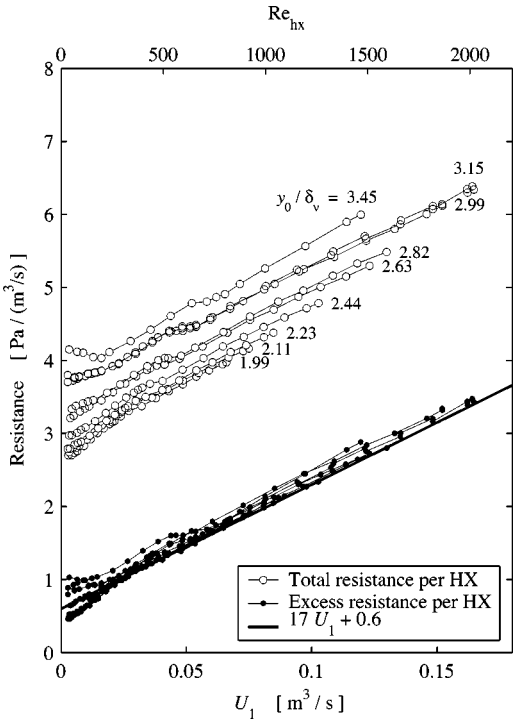


FIG. 2. Measured acoustic resistance per exchanger in the 2–6 Hz range for small exchanger separation ($2x_{hx}=5.5$ mm). Open circles are the measured resistance divided by two. Closed circles are resistance per exchanger minus the linear theory value and minus the boundary-layer value for the duct. The slope is related to the minor loss coefficient.

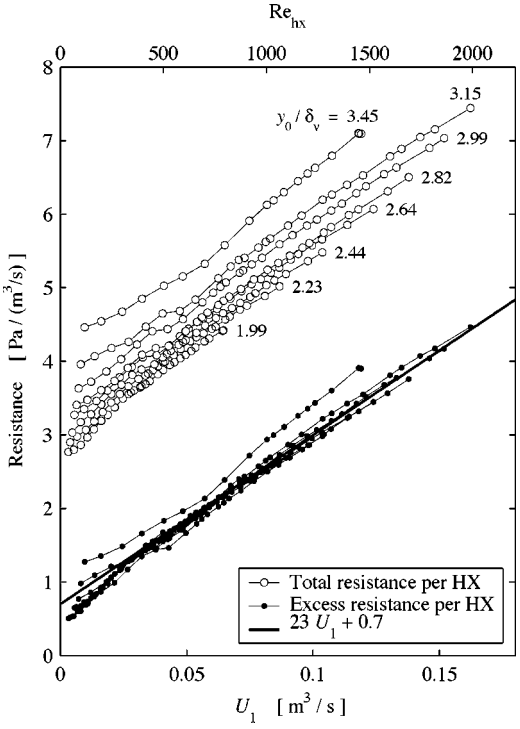


FIG. 4. Measured acoustic resistance per exchanger in the 2–6 Hz range for large exchanger separation ($2x_{hx}=54.4$ mm). Open circles are the measured resistance divided by two. Closed circles are resistance per exchanger minus the linear theory value and minus the boundary-layer value for the duct. The slope is noticeably higher than in Fig. 2.

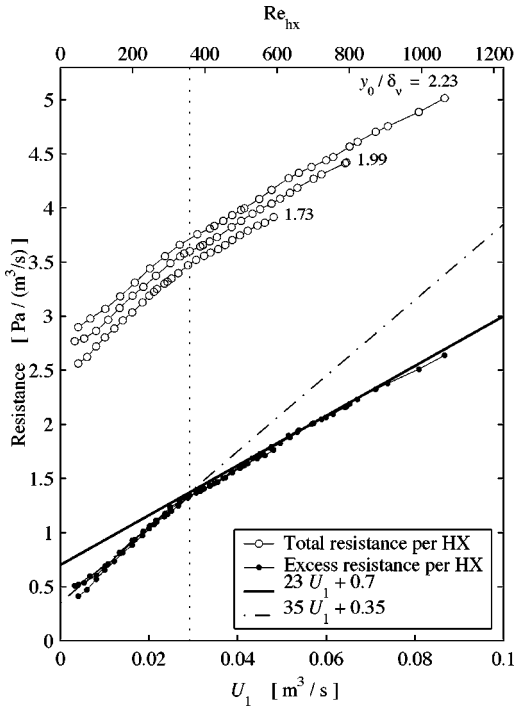


FIG. 5. Measured acoustic resistance per exchanger at 1.5, 2, and 2.5 Hz for large exchanger separation ($2x_{hx}=54.4$ mm). The knee in this data at $U_1=0.029\ \mathrm{m^3/s}$ is more sharply defined than the knee in Fig. 3.

The linear theory is well established.[4] The purpose of the measurements is to try to learn something about high-amplitude deviations from the acoustic theory. The closed circles show "excess resistance per heat exchanger," in which the theoretical linear resistances of the plates and the duct have been subtracted from the measured resistances.[6] The data collapse nearly to a single line, with a slope of 17 $\mathrm{Pa\,s^2/m^6}$. (In Figs. 2–5, all quantities are in MKS units, so the designation "$17U_1+0.6$" in the legend in Fig. 2 is shorthand notation for "$R=(17\ \mathrm{Pa\,s^2/m^6})U_1+0.6\ \mathrm{Pa\,s/m^3}$," and similarly for the legends in Figs. 3–5.)

Because the minor loss coefficient K relates pressure drop to the velocity in the heat exchanger $u_{hx}=u_1/\sigma$ according to

$$\Delta p = K\tfrac{1}{2}\rho_m u_{hx}^2, \tag{1}$$

and with

$$R\equiv\frac{\Delta p}{U_1}=\frac{\Delta p}{A_{fr}u_1}=\frac{\Delta p}{\sigma A_{fr}u_{hx}}, \tag{2}$$

an excess resistance that increases linearly with velocity corresponds to a minor loss coefficient

$$K=\frac{2\sigma^2A_{fr}^2}{\rho_m}\frac{dR}{dU_1}. \tag{3}$$

The slope of the fit line ($dR/dU_1=17\ \mathrm{Pa\,s^2/m^6}$) in Fig. 2 corresponds to an effective minor loss coefficient $K=0.125$. We say "effective" K because this is not a measurement of minor loss alone. Presumably there is also an "entrance effect," since the hydrodynamic boundary layer is almost certainly thinner near the leading edge of the exchanger. There also may be some entrance minor loss, but, for the rounded tube used here, the entrance minor loss is probably negligible. Looking at only the lowest reliable frequencies, in Fig. 3, we observe something of a "knee" at around $U_1=0.023\ \mathrm{m^3/s}$ ($\mathrm{Re_{hx}}=282$), with a higher slope at amplitudes below this value. This slope corresponds to $K=0.21$.

In the plots of large-gap data (Figs. 4 and 5), the slopes are significantly larger than those obtained from the small gap, corresponding to $K=0.17$ in the high amplitude region and $K=0.26$ at low amplitude for the lowest frequencies. The interaction between the exchangers in the small-gap case acts to lessen the overall effective minor loss coefficient. This is not surprising, since if the gap were reduced to zero, one entrance and one exit would be eliminated, cutting the per-exchanger effective minor loss in half. The wide gap results, therefore, are more representative of a single, isolated exchanger.

The uniform-velocity exit minor loss coefficient K_{min} for the parallel plate heat exchangers, which have porosity $\sigma=0.76$, is[2]

$$K_{min}=(1-\sigma)^2=0.057, \tag{4}$$

whereas the maximum coefficient K_{max}, expected for a fully developed parabolic flow profile, is[2]

$$K_{max}=\alpha-2\beta\sigma+\sigma^2=0.296, \tag{5}$$

where α and β are the "kinetic energy coefficient" and "momentum coefficient," which are 1.54 and 1.20 for parabolic flow between parallel plates. In our previous article on exit losses from nonuniform oscillating flow, we calculated effective values of α and β for laminar oscillating flow (Fig. 3 of Ref. 2).[7] Using these values we predict that the exit minor loss coefficient should range from a value of $K=0.18$ at $y_0/\delta_\nu=3.5$ ($\alpha_{eff}=1.31$, $\beta_{eff}=1.12$) up to $K=0.27$ at $y_0/\delta_\nu=2.0$ ($\alpha_{eff}=1.48$, $\beta_{eff}=1.18$). No trend in slope with frequency is discernable in Fig. 2. To the extent that the plates are short compared to the hydrodynamic entrance length, the flow is arguably not "fully developed" at the exit of the tubes, even in the oscillating flow sense that is assumed in Ref. 2. This would reduce the exit minor loss at higher amplitudes.

At low amplitudes, we expect the exit velocity profile to be close to that of the linear theory used in Ref. 2. The low-frequency, small-amplitude, wide-gap value $K=0.26$ is quite close to the value 0.27 predicted by Ref. 2, and is 4.6 times larger than the value 0.057 obtained from $(1-\sigma)^2$. This is evidence in support of the method used in Ref. 2, albeit for a single value of σ.

Shah and London[8] give the steady-flow hydrodynamic entry length L_{hy} for parallel plates as

$$\frac{L_{hy}}{4y_0}=0.031\,25+0.011\,\mathrm{Re}. \tag{6}$$

The knees in Figs. 3 and 5 occur at Reynolds numbers (for the gas *within* the exchangers) of around $\mathrm{Re_{hx}}=282$ and 355, for which Eq. (6) puts the entry lengths at 43 and 53 mm. These entry lengths are 2–2.4 times the length of an exchanger. It might have been expected, then, that the knees would have occurred at lower values of U_1. It is also curious

that the knee occurs at a higher velocity for the wide-gap data, though the location of the knee on the small-gap graph is somewhat vague, and the difference may not be significant. It should also be pointed out that, at these low frequencies, the peak gas displacement is roughly the same as the 22 mm length of the exchangers at the knee, but again, the poor definition of the knee limits our ability to comment further.

None of the plots of excess resistance (Figs. 2–5) has an intercept of zero. That is, there is some real or apparent linear resistance not accounted for by the linear theory for the plates and the ducts. Unfortunately, we do not consider the data at the very lowest amplitudes to be sufficiently reliable to judge whether this is a real or apparent offset in resistance at $U_1=0$. To understand what we mean by an ''apparent'' linear resistance, consider Fig. 3, and imagine that the low-amplitude data extrapolate along the straight line $R=28U_1+0.35$ to a value of $R=0.35\,\mathrm{Pa/(m^3/s)}$ at $U_1=0$. This 0.35 would then be a ''real'' linear resistance of $R_0=0.35\,\mathrm{Pa/(m^3/s)}$ because this would be the actual value obtained in the limit as $U_1\rightarrow 0$. The high-amplitude fit line, $R=17U_1+0.60$, on the other hand, has an intercept of 0.60 $\mathrm{Pa/(m^3/s)}$. The high-amplitude resistance, therefore, is *not* given by $R=17U_1+R_0$, but rather by $R=17U_1+R_0+R_{\mathrm{app}}$, where $R_{\mathrm{app}}=0.25\,\mathrm{Pa/(m^3/s)}$ the additional ''apparent'' linear resistance caused by the downward curvature of the excess resistance. Petculescu and Wilen[9] made very careful, accurate measurements of minor losses in jet pumps. Their data also exhibit an apparent linear resistance [see their Figs. 4(a) and 5(a)]. This means the two quantities *linear resistance* and *effective minor loss coefficient* are not sufficient to calculate the high-amplitude resistive pressure drop. A full description would have to include the apparent excess linear resistance as well.

In conclusion, these measurements support the general approach of calculating pressure drops in thermoacoustic devices by starting with the results of linear theory and adding a minor loss correction, with some caveats. The exit minor loss coefficient used in the correction may be larger than that predicted by $K=(1-\sigma)^2$, due to the development of a nonuniform velocity profile. On the other hand, for high amplitudes in very short ducts, such as these heat exchangers, the minor loss coefficient may be less than that predicted by Ref. 2 because full development of the oscillating-flow velocity profile does not occur. Other influences of ''entry effects'' have yet to be worked out.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research, the Penn State Applied Research Laboratory, and the Pennsylvania Space Grant Consortium.

[1] R. S. Wakeland and R. M. Keolian, ''Effectiveness of parallel-plate heat exchangers in thermoacoustic devices,'' J. Acoust. Soc. Am. (in press).
[2] R. S. Wakeland and R. M. Keolian, ''Influence of velocity profile nonuniformity on minor losses for flow exiting thermoacoustic heat exchangers,'' J. Acoust. Soc. Am. **112**, 1249–1252 (2002).
[3] R. S. Wakeland and R. M. Keolian, ''Measurements of resistance of individual square-mesh screens to oscillating flow at low and intermediate Reynolds numbers,'' J. Fluids Eng. **125**, 851–862 (2003).
[4] G. Swift, *Thermoacoustics: A Unifying Perspective for Some Engines and Refrigerators* (Acoustical Society of America, Mellville, NY, 2002).
[5] R. S. Wakeland, ''Heat exchangers in oscillating flow, with application to thermoacoustic devices that have neither stack nor regenerator,'' Ph.D. thesis, The Pennsylvania State University, 2003.
[6] The linear resistances are given by $R=(\omega\rho_m L/\sigma A_{\mathrm{fr}})\mathrm{Imag}(-1/1-f_\nu)$, where L is the length of the element. The ''Rott functions'' f_ν are given by Eqs. (4.56) and (4.58) of Ref. 4 for the duct and heat exchangers, respectively. For the exchangers, $\sigma=0.76$ and $L=22$ mm; for the duct, $\sigma=1$ and L is the total length of open duct between the pressure sensors, which depends on the gap size.
[7] There is a typo in the labeling of Fig. 3 of Ref. 2. The tick marker ''1.000'' on the β_{eff} scale should be ''1.100.''
[8] R. K. Shah and A. L. London, *Laminar Flow Forced Convection in Ducts, Advances in Heat Transfer*, supplement 1, edited by T. F. Irvine, Jr. and J. P. Harnett, Eq. (291) (Academic, New York, 1978).
[9] A. Petculescu and L. A. Wilen, ''Oscillatory flow in jet pumps: Nonlinear effects and minor losses,'' J. Acoust. Soc. Am. **113**, 1282–1292 (2003).

A closed form solution for the mobility of an edge-excited, semi-infinite plate

J. X. Su
Acoustics Research Unit, University of Liverpool, Liverpool L69 3BX, United Kingdom

A. T. Moorhouse[a)]
Acoustics Research Centre, University of Salford, Salford M5 4WT, United Kingdom

(Received 20 February 2002; revised 7 February 2004; accepted 10 February 2004)

Closed form analytical solutions are given for the point force and point moment mobilities at the free edge of a semi-infinite plate. Previously, these solutions have been known only in terms of unsolved integrals. The solutions give mobility as functions of Poisson's ratio times the characteristic mobility. The edge-excited plate therefore can be added to the relatively small number of structures for which exact solutions are available. The value of such solutions lies in allowing computer simulations, and testing for approximate methods. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1698861]

PACS numbers: 43.40.Dx, 43.40.At [ANN] Pages: 2075–2082

I. INTRODUCTION

The problem of edge-excited plates was first addressed by Eichler in 1964.[1] He presented solutions for the semi-infinite plate in the form of integrals, which needed to be evaluated numerically. The same problem was also investigated by Mogilevsky[2] in 1993 for isentropic plates, and again the solutions were given in terms of unsolved integrals. More recently, Gunda *et al.*[3] presented an analysis of semi-infinite plate Green's functions using the method of images. Again, numerical evaluation of integrals was required. At about the same time, Kauffmann[4] applied Eichler's analysis to evaluation of power absorbed by the plate, and made a more accurate numerical evaluation of the integrals than was possible in the 1960s. On the basis of the numerical results, he made a number of conjectures, namely (a) that the point mobility is pure real, (b) that the real parts of the normal and tangential moment mobilities are equal, and (c) that the cross mobility has a phase of $\pi/4$. The objective of this paper is to evaluate some of the integrals so as to yield closed form expressions for diagonal elements of the mobility matrix, and secondly to prove Kauffmann's conjectures.

II. DIFFERENTIAL EQUATION AND BOUNDARY CONDITION

For a thin elastic plate of constant thickness h, the transverse motion, w, under sinusoidal excitation of angular frequency ω satisfies the following differential equation:

$$(D\nabla^4-\omega^2 m'')w=0, \tag{1}$$

where $D=Eh^3/12(1-\nu^2)$ is the bending stiffness, E is Young's modulus, ν is Poisson's ratio, and m'' is the mass per unit area. The notation used throughout this paper is consistent with both Eichler and Kauffmann.

Consider a semi-infinite plate (see Fig. 1). The plate occupies $-\infty<y<\infty$ $x\leqslant 0$ with a free edge at $x=0$ except for the area $|y|<a$ where it is excited by an arbitrary load. The transverse force, normal and tangential moments at the edge are related to the plate displacement by the boundary conditions

$$-D\frac{\partial}{\partial x}\left[\frac{\partial^2}{\partial x^2}w+(2-\nu)\frac{\partial^2}{\partial y^2}w\right]=N-\frac{\partial H}{\partial y}=f(y) \tag{2}$$

$$-D\left[\frac{\partial^2}{\partial x^2}w+\nu\frac{\partial^2}{\partial y^2}w\right]=m(y) \tag{3}$$

where N is the distribution of the transverse force per unit length, H is the distribution of the normal moment per unit length, $f(y)$ is the externally applied force per unit length, and $m(y)$ is the externally applied tangential moment per unit length.

III. POINT MOBILITY MATRIX

The point mobilities at $x=0$ and $y=0$ are of interest, and can be obtained by letting $f(y)=N-\partial H/\partial y=F_z\delta(y)-M_n\delta'(y)$ and $m(y)=M_t\delta(y)$ and solving for the velocity and angular velocity at $x=0$ $y=0$,

$$\begin{bmatrix}\Omega_n\\ \Omega_t\\ V_z\end{bmatrix}=i\omega\begin{bmatrix}\partial w/\partial y\\ -\partial w/\partial x\\ w\end{bmatrix}=[Y]\begin{bmatrix}M_n\\ M_t\\ F_z\end{bmatrix}, \tag{4}$$

where Ω_n is the normal angular velocity, Ω_t is the tangential angular velocity, V_z is the transverse velocity, M_n, M_t are the normal and tangential moments, and F_z is the transverse force. Y is the point mobility matrix and can be expressed as

$$Y=(m''D)^{-1/2}\begin{bmatrix}y_{\Omega_n M_n}k^2 & 0 & 0\\ 0 & y_{\Omega_t M_t}k^2 & y_{\Omega_t F_z}k\\ 0 & y_{V_z M_t}k & y_{V_z F_z}\end{bmatrix}, \tag{5}$$

[a)]Author to whom correspondence should be addressed. Electronic mail: a.t.moorhouse@salford.ac.uk

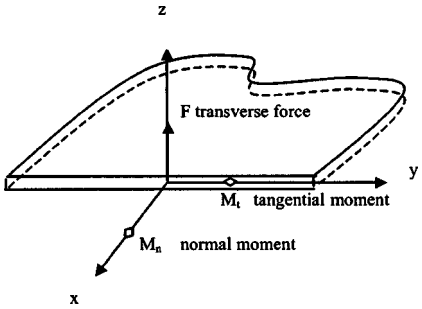


FIG. 1. Semi-infinite plate.

where $k=m''\omega^2/D$ is the free wavenumber. The elements of the matrix are expressed in terms of a nondimensional wavenumber p (where p is the wavenumber nomalized by dividing by the free wavenumber) as

$$\mathrm{Re}(y_{\Omega_n M_n})=\frac{2}{\pi}\int_0^{-1} p^2\mu_1(p)M_+^2(p)\frac{dp}{\mathrm{Det}(p)}+2p_1^2\frac{1}{\gamma}, \tag{6}$$

$$\mathrm{Im}(y_{\Omega_n M_n})=\frac{2}{\pi}PV\int_1^{\infty}\frac{p^2dp}{\det(p)} -\frac{2}{\pi}\int_0^1\frac{\mu_+(p)M_-^2(p)p^2\,dp}{\mathrm{Det}(p)}, \tag{7}$$

$$\mathrm{Re}(y_{\Omega_t M_t})=\frac{2}{\pi}\int_0^1\frac{\mu_1(p)\mu_+^2(p)M_-^2(p)dp}{\mathrm{Det}(p)} +\frac{2\mu_-(p_1)\mu_+(p_1)}{\gamma}, \tag{8}$$

$$\mathrm{Im}(y_{\Omega_t M_t})=\frac{2}{\pi}\left(\int_0^1\frac{\mu_+(p)\mu_1^2(p)M_+^2(p)dp}{\mathrm{Det}(p)} +PV\int_1^{\infty}\frac{\mu_-(p)\mu_+(p)dp}{\det(p)}\right), \tag{9}$$

$$\mathrm{Re}(y_{\Omega_t F_z})=-\frac{2}{\pi}\int_0^1\mu_1(p)\mu_+(p)M_-(p)M_+(p) \times\frac{dp}{\mathrm{Det}(p)}-[\mu_-(p_1)M_+(p_1) +\mu_+(p_1)M_-(p_1)]\frac{1}{\gamma}, \tag{10}$$

$$\mathrm{Im}(y_{\Omega_t F_z})=-\frac{1}{\pi}\int_0^1[\mu_1^2(p)M_+^3(p) -\mu_+^2(p)M_-^3(p)]\frac{dp}{\mathrm{Det}(p)} -\frac{1}{\pi}PV\int_1^{\infty}\mu_+(p)M_-(p) +\mu_-(p)M_+(p)\frac{dp}{\det(p)}, \tag{11}$$

$$\mathrm{Re}(y_{V_z F_z})=\frac{2}{\pi}\int_0^1\mu_1(p)M_+^2(p)\frac{dp}{\mathrm{Det}(p)}+2\frac{1}{\gamma}, \tag{12}$$

$$\mathrm{Im}(y_{V_z F_z})=\frac{2}{\pi}PV\int_1^{\infty}\frac{dp}{\det(p)}-\frac{2}{\pi}\int_0^1\mu_+M_-^2(p)\frac{dp}{\mathrm{Det}(p)}, \tag{13}$$

where $\mu_1(p)\equiv\sqrt{1-p^2}$, $\mu_\pm(p)\equiv\sqrt{p^2\pm1}$. p should be understood as a real integration variable. To be clear, here and from now on the authors distinguish the square root symbol in the real field from square root symbols in the complex field. "$\sqrt{\ }$" indicates a square root in the real field, the function $\sqrt{x}$ should therefore be understood as a single-valued function, and x must be real and positive. "$(\ )^{1/2}$" indicates a square root in the complex field. The function $(w)^{1/2}$ is multiple-valued, and a value can be determined after choosing a single-value branch. For example the square root of "-3" is written as $(-3)^{1/2}$ not -3. In this paper $\sqrt{-3}$ is meaningless. The value of the function $(-3)^{1/2}=\sqrt{3}i$ or $(-3)^{1/2}=-\sqrt{3}i$ depends on how a single-value branch is chosen.

In Eqs. (6)–(13)

$$\begin{aligned} &M_\pm(p)\equiv1\pm(1-\nu)p^2,\\ &\det(p)=\mu_-(p)M_+^2(p)-\mu_+(p)M_-^2(p), \end{aligned} \tag{14}$$

$$\begin{aligned} \mathrm{Det}(p)&=(1-p^2)M_+^4(p)+(1+p^2)M_-^4(p)\\ &=2[-(3+\nu)(1-\nu)^3p^8\\ &\quad+2(1-3\nu)(1-\nu)p^4+1]\\ &=-2(3+\nu)(1-\nu)^3(p^4-p_1^4)(p^4+p_2^4), \end{aligned} \tag{15}$$

$$p_1=\left|(1-3\nu+2\sqrt{\nu^2+(1-\nu)^2})/(1-\nu)^2(3+\nu)\right|^{1/4}, \tag{16}$$

$$p_2=\left|(3\nu+2\sqrt{\nu^2+(1-\nu)^2}-1)/(1-\nu)^2(3+\nu)\right|^{1/4}. \tag{17}$$

p_1 is one root of $\det(p)$ on the real axis and lies between 1 and $[4(2\sqrt{2}-1)/7]^{1/4}$ for $0\leqslant\nu\leqslant0.5$.[2] The integration variable, p, should be understood as a real integration variable and PV stands for principle value in Eqs. (7), (9), (11), (13) (there is one simple pole at $p=p_1$ between land ∞ on the real axis). $1/\gamma$ is the residue of the function $1/\det(p)$ at the simple pole $p=p_1$. It can be expressed as

$$\frac{1}{\gamma}=\lim_{p\to p_1}\frac{p-p_1}{\det(p)}=\frac{\mu_-(p_1)M_+^2(p_1)+\mu_+(p_1)M_-^2(p_1)}{8(3+\nu)(1-\nu)^3p_1^3(p_1^4+p_2^4)}. \tag{18}$$

The second form on the right-hand side of Eq. (18) differs from Kauffmann's (although it is still equivalent) and is used in the following proofs. Expressions (6)–(13) can be obtained directly from Eqs. (47)–(52) of Kaufmann's paper[4] or from Eqs. (18)–(25) of Eichler's paper[1] in the case of $ka\to0$, but are repeated here for completeness.

Eichler and Kaufmann both evaluated numerical results for the point mobility but Kaufmann's, with greater computing power available, are more accurate.[4] From his numerical

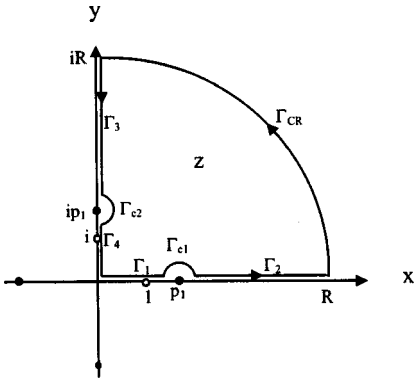


FIG. 2. Contour of integration for evaluation of the imaginary part of the point force mobility.

results for Poisson's ratio $\nu=0.3$, he conjectured that the following relations might be satisfied for all values of Poisson's ratio $(0<\nu<0.5)$:

$$\mathrm{Im}(y_{V_zF_z})=0, \quad \mathrm{Re}(y_{\Omega_tF_z})=\mathrm{Im}(y_{\Omega_tF_z}),$$

$$\mathrm{Re}(y_{\Omega_nM_n})=\mathrm{Re}(y_{\Omega_tM_t})$$

These relations are formally proved in the following section.

IV. PROOF FOR THE CONJECTURES AND ANALYTIC EXPRESSIONS FOR POINT MOBILITY

It will now be shown that the above conjectures can be proved and most elements of the point mobility matrix can be expressed by closed form expressions instead of integration formulas. This is the main aim of this paper.

Some definitions of symbols are needed before completion of the proof.

$\int_\Gamma F(z)dz$ stands for integration on the complex plane z, the real and imaginary axes of which are denoted x and y. $z\equiv x+yi$ is a complex integration variable, and x and y real integration variables. Γ is the contour of integration on the complex plane z.

A. Proof of $\mathrm{Im}(y_{V_zF_z})=0$

To prove that Eq. (13) is equal to zero we consider the following integral around a closed contour Γ:

$$\int_\Gamma \frac{dz}{\widetilde{\det}(z)}, \quad \widetilde{\det}(z)\equiv\tilde{\mu}_-(z)M_+^2(z)-\tilde{\mu}_+(z)M_-^2(z), \tag{19}$$

where $M_\pm(z)\equiv 1\pm(1-\nu)z^2$ and $\tilde{\mu}_\pm(z)$ is a single-valued branch formed by the square root with a positive real part of the multiple-valued function $(z^2\pm 1)^{1/2}$.[5] The contour of integration Γ is shown in Fig. 2 in which Γ_{c1} and Γ_{c2} are semi-circles of radius ε. The integrand is analytic within Γ, so the integral should be zero, i.e.,

$$\int_\Gamma \frac{dz}{\widetilde{\det}(z)}=0, \tag{20}$$

and this integration can be rewritten as

$$\int_\Gamma \frac{1}{\widetilde{\det}(z)}dz$$

$$=\int_{\Gamma_1}+\int_{\Gamma_{c1}}+\int_{\Gamma_2}+\int_{\Gamma_{CR}}+\int_{\Gamma_3}+\int_{\Gamma_{c2}}+\int_{\Gamma_4}\frac{1}{\widetilde{\det}(z)}dz$$

$$=I_1+I_{c1}+I_2+I_{CR}+I_3+I_{c2}+I_4. \tag{21}$$

The subintegrals in (21) are now evaluated. To evaluate I_1, we rewrite in terms of $\mathrm{Det}(x)$:

$$I_1=i\int_0^1\frac{\mu_1(x)M_+^2(x)dx}{-\mathrm{Det}(x)}+\int_0^1\frac{\mu_+(x)M_-^2(x)dx}{-\mathrm{Det}(x)}+\int_1^{p_1-\varepsilon}\frac{dx}{\det(x)}, \tag{22}$$

$$\mu_1(x)\equiv\sqrt{1-x^2},\mu_\pm(x)\equiv\sqrt{x^2\pm 1},$$

$$M_\pm(x)=1\pm(1-\nu)x^2 \quad \text{and}$$

$$\det(x)=\mu_-(x)M_+^2(x)-\mu_+(x)M_-^2(x), \tag{23}$$

$$I_2=\int_{p_1+\varepsilon}^R\frac{dx}{\det(x)}.$$

To evaluate I_3 let iy replace z so that we have

$$I_3=\int_{iR}^{i(p_1+\varepsilon)}\frac{dz}{\widetilde{\det}(z)}=\int_{p_1+\varepsilon}^R\frac{dy}{\det(y)}=I_2, \tag{24}$$

$$I_4=\int_{i(p_1-\varepsilon)}^0\frac{dz}{\widetilde{\det}(z)}=i\int_0^1\frac{\mu_1(y)M_+^2(y)dy}{\mathrm{Det}(y)}+\int_0^1\frac{\mu_+(y)M_-^2(y)dy}{-\mathrm{Det}(y)}+\int_1^{p_1-\varepsilon}\frac{dy}{\det(y)}, \tag{25}$$

while

$$\lim_{\varepsilon\to 0}(I_{c1}+I_{c2})=-\pi i\left(\mathrm{Res}\frac{1}{\widetilde{\det}(p_1)}+\mathrm{Res}\frac{1}{\widetilde{\det}(ip_1)}\right)=0, \tag{26}$$

$$\lim_{R\to\infty}I_{CR}=0, \tag{27}$$

because if

$$R>\max\left(1,1/(1-\nu),\left(\frac{3|2(1-3\nu)(1-\nu)|}{(3+\nu)(1-\nu)^3}\right)^{1/4},\left(\frac{1}{(3+\nu)(1-\nu)^3}\right)^{1/8}\right) \tag{28}$$

$$\left|\int_{\Gamma_{CR}}\frac{1}{\widetilde{\det}(z)}\right|\leqslant\frac{24\sqrt{2}}{2|(3+\nu)(1-\nu)^3||R|^2}\frac{\pi R}{2R},$$

so adding the nonzero integrals, I_1, I_2, I_3, I_4 some terms cancel leaving

$$\lim_{\substack{R\to\infty\\ \varepsilon\to 0}}\int_\Gamma\frac{1}{\widetilde{\det}(z)}dz=2\left(PV\int_1^\infty\frac{dx}{\det(x)}-\int_0^1\frac{\mu_+(x)M_-^2(x)dx}{\mathrm{Det}(x)}\right)$$

$$=0. \tag{29}$$

Comparing Eqs. (29) and (13) shows that the imaginary part of the point mobility is identically zero, in other words that edge mobilities, like the point mobility of an infinite plate, are pure real.

B. Proof of $\mathrm{Re}(y_{\Omega_t F_z})=\mathrm{Im}(y_{\Omega_t F_z})$

The proof follows similar lines to the above and proceeds by considering the integral:

$$\int_\Gamma F(z)dz=\int_\Gamma \frac{\tilde{\mu}_-(z)M_+(z)+\tilde{\mu}_+(z)M_-(z)dz}{\tilde{\det}(z)}. \quad (30)$$

The single-valued branch and contour of integration are selected as above. Again, according to the residue theorem, this integral is zero around the contour and can be expressed as a sum of integrals (denoted by the same symbols as above for simplicity), i.e.,

$$0=\int_{\Gamma_1}+\int_{\Gamma_{c1}}+\int_{\Gamma_2}+\int_{\Gamma_{CR}}+\int_{\Gamma_3}+\int_{\Gamma_{c2}}+\int_{\Gamma_4}F(z)dz$$
$$=I_1+I_{c1}+I_2+I_{CR}+I_3+I_{c2}+I_4. \quad (31)$$

By rewriting the integrand of (30) in terms of $\mathrm{Det}(z)$ we get

$$I_1=\int_0^1 \frac{\mu_1^2(x)M_+^3(x)-\mu_+^2(x)M_-^3(x)}{\mathrm{Det}(x)}dx$$
$$-i\int_0^1 \frac{2\mu_1(x)M_+(x)\mu_+(x)M_-(x)}{\mathrm{Det}(x)}dx$$
$$+\int_1^{p_1-\varepsilon} \frac{\mu_-(x)M_+(x)+\mu_+(x)M_-(x)}{\det(x)}dx, \quad (32)$$

$$I_2=\int_{p_1+\varepsilon}^R \frac{\mu_-(x)M_+(x)+\mu_+(x)M_-(x)}{\det(x)}dx. \quad (33)$$

To evaluate I_3 let iy replace z giving

$$I_3=i\int_{p_1+\varepsilon}^R \frac{\mu_-(y)M_+(y)+\mu_+(y)M_-(y)}{\det(y)}dy=iI_2, \quad (34)$$

$$I_4=-\int_0^1 \frac{2\mu_1(y)M_+(y)+\mu_+(y)M_-(y)}{\mathrm{Det}(y)}dy$$
$$i\int_0^1 \frac{\mu_1^2(y)M_+^3(y)-\mu_+^2(y)M_-^3(y)}{\mathrm{Det}(y)}dy$$
$$+i\int_1^{p_1-\varepsilon} \frac{\mu_-(y)M_+(y)+\mu_+(y)M_-(y)}{\det(y)}dy, \quad (35)$$

while

$$\lim_{R\to\infty} I_{\Gamma_{CR}}=0, \quad (36)$$

$$\lim_{\varepsilon\to 0} I_{\Gamma_{c1}}=-\pi i\frac{\mu_-(p_1)M_+(p_1)+\mu_+(p_1)M_-(p_1)}{\gamma}, \quad (37)$$

$$\lim_{\varepsilon\to 0} I_{\Gamma_{c2}}=-\pi\frac{\mu_-(p_1)M_+(p_1)+\mu_+(p_1)M_-(p_1)}{\gamma}, \quad (38)$$

so

$$\lim_{\substack{R\to\infty\\ \varepsilon\to 0}}\int_\Gamma F(z)dz=(1+i)\int_0^1 \frac{\mu_1^2(x)M_+^3(x)-\mu_+^2(x)M_-^3(x)-2\mu_1(x)M_+(x)\mu_+(x)M_-(x)}{\mathrm{Det}(x)}dx+(1+i)PV$$
$$\times\int_1^\infty \frac{\mu_-(x)M_+(x)+\mu_+(x)M_-(x)}{\det(x)}dx-\pi(1+i)[\mu_+(p_1)M_-(p_1)+\mu_-(p_1)M_+(p_1)]\frac{1}{\gamma}=0. \quad (39)$$

Multiplying both sides of Eq. (39) by factor $1/\pi(1+i)$ shows that the difference of Eqs. (10) and (11) is zero, thereby proving that the phase of the cross mobility is constant at $\pi/4$ irrespective of the Poisson's ratio.

C. Real part of point force mobility

To evaluate the real part of the point mobility we proceed by considering the integral

$$\int_\Gamma F(z)dz=\int_\Gamma \frac{\tilde{\mu}_-(z)f_1(z)dz}{\mathrm{Det}(z)},$$
$$f_1(z)\equiv(1+(1-\nu)z^2)^2, \quad (40)$$

where $\tilde{\mu}_-(z)$ is a single-valued branch formed by the square root with a positive **imaginary part** of the multiple-valued function $(z^2-1)^{1/2}$.

The contour of integration Γ is shown in Fig. 3. The integrand is analytic within Γ except at three simple poles: ($z_1=ip_1$ $z_2=p_2e^{\pi i/4}$ $z_3=p_2e^{3\pi i/4}$), where p_2 is defined in Eqs. (17).

Expression (40) can be rewritten as

$$\int_\Gamma F(z)dz=\int_{\Gamma_1}+\int_{\Gamma_{c2}}+\int_{\Gamma_2}+\int_{\Gamma_3}+\int_{\Gamma_{c1}}+\int_{\Gamma_4}$$
$$+\int_{\Gamma_{CR}}F(z)dz$$
$$=I_1+I_{c2}+I_2+I_3+I_{c1}+I_4+I_{CR}, \quad (41)$$

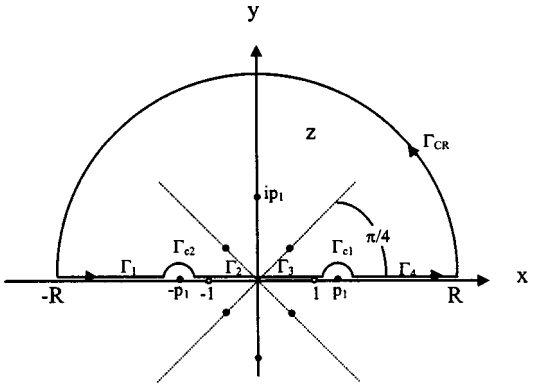


FIG. 3. Contour of integration for evaluation of the real part of the point force mobility.

$$I_1=\int_{-R}^{-(p_1+\varepsilon)}\frac{-\mu_-(x)M_+^2(x)}{\mathrm{Det}(x)}dx$$

$$=\int_{p_1+\varepsilon}^{R}\frac{-\mu_-(x)M_+^2(x)}{\mathrm{Det}(x)}dx \tag{42}$$

$$I_2=\int_{-p_1+\varepsilon}^{-1}\frac{-\mu_-(x)M_+^2(x)}{\mathrm{Det}(x)}dx+i\int_{-1}^{0}\frac{\mu_1(x)M_+(x)}{\mathrm{Det}(x)}dx$$

$$=i\int_0^1\frac{\mu_1(x)M_+(x)}{\mathrm{Det}(x)}dx-\int_1^{p_1-\varepsilon}\frac{\mu_-(x)M_+^2(x)}{\mathrm{Det}(x)}dx, \tag{43}$$

$$I_3=i\int_0^1\frac{\mu_1(x)M_+(x)}{\mathrm{Det}(x)}dx+\int_1^{p_1-\varepsilon}\frac{\mu_-(x)M_+^2(x)}{\mathrm{Det}(x)}dx$$

$$=-I_2^*, \tag{44}$$

where "*" means complex conjugate,

$$I_4=\int_{p_1+\varepsilon}^{R}\frac{\mu_-(x)M_+^2(x)}{\mathrm{Det}(x)}dx=-I_1, \tag{45}$$

while

$$\lim_{\varepsilon\to 0} I_{c2}=\pi i\mu_-(p_1)M_+^2(p_1)/4\Lambda_1 \tag{46}$$

$$\Lambda_1\equiv 2(3+\nu)(1-\nu)^3p_1^3(p_1^4+p_2^4), \tag{47}$$

$$\lim_{\varepsilon\to 0} I_{c1}=\pi i\mu_-(p_1)M_+^2(p_1)/4\Lambda_1, \tag{48}$$

$$\lim_{R\to\infty} I_{CR}=0. \tag{49}$$

According to the residue theorem the value of the integral is given by the sum of the residues of the poles:

$$\int_\Gamma F(z)dz=2\pi i(\mathrm{Res}F(z_1)+\mathrm{Res}F(z_2)+\mathrm{Res}F(z_3)), \tag{50}$$

$$\mathrm{Res}F(z_1)=\mu_+(p_1)M_-^2(p_1)/4\Lambda_1, \tag{51}$$

$$\mathrm{Res}F(z_2)=(\alpha+\beta i)(1+i(1-\nu)p_2^2)^2e^{-\pi i3/4})/4\Lambda_2, \tag{52}$$

$$\Lambda_2\equiv 2(3+\nu)(1-\nu)^3p_2^3(p_1^4+p_2^4), \tag{53}$$

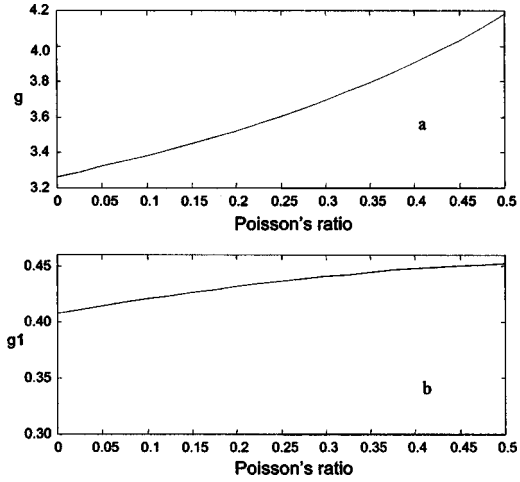


FIG. 4. Ratio of point force mobility to (a) infinite plate mobility and (b) Y_0.

where $\alpha=(\sqrt{\sqrt{1+p_2^4}-1})\sqrt{2}/2$, $\beta=(\sqrt{\sqrt{1+p_2^4}+1})\sqrt{2}/2$ are the real and imaginary parts of $(p_2^2i-1)^{1/2}$, respectively, and

$$\mathrm{Res}F(z_3)=(-(-\alpha+\beta i)(1-i(1-\nu)p_2^2)^2e^{\pi i3/4})/4\Lambda_2. \tag{54}$$

Note that $\mathrm{Res}F(z_2)$ is the complex conjugate of $\mathrm{Res}F(z_3)$.

Finally, we obtain

$$\frac{2}{\pi}\int_0^1\frac{\mu_1(x)M_+^2(x)}{\mathrm{Det}(x)}dx$$

$$=\frac{\sqrt{2}}{2}\mathrm{Re}((\beta-\alpha+(\beta+\alpha)i)f_1(p_2e^{3\pi i/4})/\Lambda_2)$$

$$+(\mu_+(p_1)f_1(ip_1)-\mu_-(p_1)f_1(p_1)/2\Lambda_1), \tag{55}$$

where f_1 is defined in Eqs. (40)

The second term of Eq. (55) is zero since p_1 is a root of $\det(p)$, so that substitution of Eq. (55) into Eq. (12) yields

$$\mathrm{Re}(y_{V_zF_z})=\frac{\sqrt{2}}{2}\mathrm{Re}((\beta-\alpha+(\beta+\alpha)i)f_1(p_2e^{3\pi i/4})/\Lambda_2)$$

$$+\frac{2}{\gamma}. \tag{56}$$

Thus, the real part of the force mobility at an edge varies from the infinite plate value by a function which depends only on Poisson's ratio, and which can be evaluated from Eq. (56), using expressions (16), (17), (40), (53), and the definitions of α and β given below Eq. (53) (further simplification is not possible). The function is discussed later and plotted in Fig. 4.

D. Real part of the point normal moment mobility

First we must rewrite the first term in Eq. (6):

$$\int_0^1 \frac{x^2\mu_1(x)M_+^2(x)dx}{\mathrm{Det}(x)}$$

$$= \int_0^1 \frac{(1-\nu)^2\mu_1(x)x^2\,dx}{-2(3+\nu)(1-\nu)^3(x^4-p_1^4)}$$

$$+ \int_0^1 \frac{\mu_1(x)[(1-(1-\nu)^2p_2^4)x^2+2(1-\nu)x^4]dx}{\mathrm{Det}(x)}. \quad (57)$$

The first term on the right-hand side of Eq. (57) can be obtained directly:

$$\int_0^1 \frac{(1-\nu)^2\sqrt{1-x^2}x^2\,dx}{-2(3+\nu)(1-\nu)^3(x^4-p_1^4)}$$

$$= \frac{(1-\nu)^2}{4(3+\nu)(1-\nu)^3}\left(2-\frac{\mu_-(p_1)}{p_1}-\frac{\mu_+(p_1)}{p_1}\right)\frac{\pi}{2}. \quad (58)$$

The second term can be obtained using the same method as described above for the real part of the force mobility and is given by

$$\frac{\sqrt{2}}{2}\mathrm{Re}[(\beta-\alpha+(\beta+\alpha)i)f_2(p_2e^{3\pi i/4})/\Lambda_2]$$

$$+[\mu_+(p_1)f_2(ip_1)-\mu_-(p_1)f_2(p_1)]/2\Lambda_1, \quad (59)$$

where

$$f_2(z)\equiv[1-(1-\nu)^2p_2^4]z^2+2(1-\nu)z^4. \quad (60)$$

It is noted that p_2 satisfies the following expression:

$$[-(3+\nu)(1-\nu)^3(-p_2^4)^2+2((1-3\nu)(1-\nu)(-p_2^4)+1)]$$

$$=0 \quad [\text{see Eq. (15)}]. \quad (61)$$

From Eq. (61) we have one of the following two relationships:

$$[1-(1-\nu)^2p_2^4](\beta+\alpha)=2(1-\nu)p_2^2(\beta-\alpha), \quad (62)$$

$$[1-(1-\nu)^2p_2^4](\beta-\alpha)=-2(1-\nu)p^2(\alpha+\beta). \quad (63)$$

For the case $0\leqslant\nu\leqslant 0.5\,(1-(1-\nu)^2p_2^4)>0$, while $2(1-\nu)p_2^2>0$ and $\alpha+\beta>0$, $\beta-\alpha>0$ so only relationship (62) is correct. Hence, the first term of Eq. (59) is zero. The remaining second term of Eq. (59) together with Eq. (58) is substituted into (57) and the resulting expression into Eq. (6) to give

$$\mathrm{Re}(y_{\Omega_m M_n})=\frac{\mu_+(p_1)f_2(ip_1)-\mu_-(p_1)f_2(p_1)}{2\Lambda_1}$$

$$+\frac{(1-\nu)^2}{4(3+\nu)(1-\nu)^3}\left(2-\frac{\mu_-(p_1)}{p_1}-\frac{\mu_+(p_1)}{p_1}\right)+2p_1^2\frac{1}{\gamma}. \quad (64)$$

Again, the mobility at an edge varies from the infinite plate value by a closed form function of Poisson's ratio [see Eqs. (16), (18), (47), (60) and the definition of μ_+, μ_- below Eq. (13) for definitions of symbols as functions of ν which cannot be simplified further].

E. Proof of $\mathrm{Re}(y_{\Omega_n M_n})=\mathrm{Re}(y_{\Omega_t M_t})$

We proceed by taking the difference between Eqs. (6) and (8):

$$\mathrm{Re}(y_{\Omega_t M_t})-\mathrm{Re}(y_{\Omega_n M_n})$$

$$=\frac{2}{\pi}\int_0^1 \frac{\mu_1(x)(M_-^2(x)-4(1-\nu)x^4)dx}{\mathrm{Det}(x)}$$

$$+2\mu_-(p_1)\mu_+(p_1)\frac{1}{\gamma}-2p_1^2\frac{1}{\gamma}. \quad (65)$$

The first term of Eq. (65) is

$$\frac{\sqrt{2}}{2}\mathrm{Re}[(\beta-\alpha+(\beta+\alpha)i)f_3(p_2e^{3\pi i/4})/\Lambda_2]$$

$$+[\mu_+(p_1)f_3(ip_1)-\mu_-(p_1)f_3(p_1)]/2\Lambda_1, \quad (66)$$

where

$$f_3(z)=[1-(1-\nu)z^2]^2-4(1-\nu)z^4. \quad (67)$$

Inserting (67) into (66) then into (65) and using the formulation (18) for $1/\gamma$, only one term remains:

$$\frac{\sqrt{2}}{2}\mathrm{Re}((\beta-\alpha)+(\beta+\alpha)i)f_3(p_2e^{3\pi i/4})/\Lambda_2. \quad (68)$$

We now prove that this remaining term is zero. Evaluating Eq. (68) gives

$$\frac{\sqrt{2}}{2}\mathrm{Re}(((\beta-\alpha)+(\beta+\alpha)i)(1+i(1-\nu)p_2^2)^2$$

$$+4(1-\nu)p_2^4)/\Lambda_2$$

$$=\frac{\sqrt{2}}{2}(\beta-\alpha)\Bigg[1-(1-\nu)^2p_2^4$$

$$-2\frac{(\beta+\alpha)}{(\beta-\alpha)}(1-\nu)p_2^2+4(1-\nu)p_2^4\Bigg]\Bigg/\Lambda_2. \quad (69)$$

Using Eq. (62) and rearranging, the numerator of Eq. (69) is rewritten

$$\frac{\sqrt{2}}{2}\frac{(\beta-\alpha)}{1-(1-\nu)^2p_2^4}[-(3+\nu)(1-\nu)^3(-p_2^4)^2$$

$$+2((1-3\nu)(1-\nu)(-p_2^4)+1)]=0. \quad (70)$$

Note that Eq. (61) is used again. The result follows:

$$\mathrm{Re}(y_{\Omega_n M_n})=\mathrm{Re}(y_{\Omega_t M_t}). \quad (71)$$

This proves that the real parts of the normal and tangential moment mobility are equal. It should be noted that the imaginary parts of the normal and tangential mobility are not identical (proof not given), even though they both numerically tend to infinity.

For the case of Poisson's ratio $\nu=0.3$ the mobility matrix is given in the following numerical form:

$$Y=(m''D)^{-1/2}\begin{bmatrix}(0.216\,45+i\infty)k^2 & 0 & 0\\ 0 & (0.216\,45+i\infty)k^2 & -0.291\,49(1+i)k^*\\ 0 & -0.291\,49(1+i)k & 0.461\,98\end{bmatrix}$$

[*numerical solution of integration expression (10)].

V. COMPARISON OF EXACT VALUES WITH PUBLISHED NUMERICAL VALUES

Point mobilities have been evaluated from Eqs. (29), (56), (64), and (71), and are compared with Kauffmann's numerical values in Table I. It is seen that Kauffmann's numerical results agree closely with the exact values. According to Eq. (5) the force point mobility was given by

$$Y_{VF}=8Y_{VF}^{\infty}y_{V_zF_z}=\sqrt{12}(m''Eh^3)^{-1/2}\sqrt{1-\nu^2}y_{V_zF_z}. \quad (72)$$

Figure 4(a) shows the edge mobility normalized by the infinite plate mobility $g=Y_{VF}/Y_{VF}^{\infty}=8y_{V_zF_z}$ versus Poisson's ratio ν. This curve corresponds to that from Fig. 2 of Ref. 4 for $\gamma_1=1$, and the shapes of both curves are almost the same. The normalized mobility is a monotonically increasing function of the Poisson's ratio ν. The range of possible values is relatively narrow with the mobility at a free edge of a plate being between 3.2 and 4.2 times higher than at an internal point far from any edges. (The corresponding factor for a beam is 4.[6])

Figure 4(b) shows the mobility normalized by $Y_0=\sqrt{12}(m''Eh^3)^{-1/2}$, i.e., $g_1=\sqrt{1-\nu^2}y_{V_zF_z}$ versus Poisson's ratio ν. It is seen to increase with the increase of Poisson's ratio ν. This is unexpected because one would expect bending stiffness to increase with Poisson's ratio ν, and hence mobility to decrease as for an infinite plate. Here the opposite is the case which is a result of using boundary conditions Eq. (2).

TABLE I. Comparison of point mobility for edge-excited semi-infinite plates, $\nu=0.3$. When $p\to\infty$, the integrand function of the first term in expression (7) and the integrand function of the second term in expression (9)$\to 1/(1-\nu)(3+\nu)p$, so the imaginary part of the point moment mobility is infinite.

	Real part	Imaginary part
	Kauffmann/this paper	Kauffmann/this paper
$y_{\Omega_n M_n}$	0.216 44/0.216 45	∞/∞
$y_{\Omega_t M_t}$	0.216 44/0.216 45	∞/∞
$y_{V_zF_z}$	0.461 96/0.461 98	$-5.97\times10^{-7}/0$

VI. INTERPRETATION OF RESULTS

It is well known that an infinite plate acts as a pure damper when excited by a point force, all energy input is propagated away and none is reflected back to the source. Section IV A proves that this is also true when a point force acts at a free edge, since the edge mobility is pure real. Since points both at the edge and far from it act as pure resistances we can conjecture that this is also true at intervening points and hence at any point on a free semi-infinite plate, although only the edge and center cases have been proven.

The results of Sec. IV C are plotted in Fig. 4(a). It shows that for $\nu=0.3$ the level of power input by a force at a free edge is a factor of 3.7 times higher than well away from the edge (compared with a factor of 4.0 for a beam). The actual ratio for a plate varies with Poisson's ratio as shown in Fig. 4(a), and as might be expected the difference between the semi-infinite and infinite cases increases with Poisson's ratio due to the stiffening effect of constrained in-plane expansions.

Section IV E confirms the conjecture that the real parts of the normal and tangential moment mobility are equal. This would be expected for an infinite plate because of circular symmetry but it is not obvious that it should be so for an edge-excited plate. It indicates that the same power is input irrespective of whether the moment acts normally or tangentially to the edge. Indeed, since there is no coupling (zero off-diagonals) between normal and tangential moments the same holds more generally for a moment applied at an arbitrary angle. In other words, a moment applied at the free edge inputs the same power irrespective of its orientation. It is interesting to compare this result with the case of an acoustic dipole, which, although not an exact analogy, bears some resemblance to a force couple and might be expected to behave in a similar way. If an acoustic dipole is brought up close to a rigid surface, it effectively becomes a quadrupole whose sound power is strongly dependent on the orientation relative to the surface.[6] The same is true for a pressure release boundary. By simple analogy, the orientation of the moment at the free edge of a plate might be expected to influence power input: the finding that it has no effect therefore at first appears counter-intuitive. The apparent contradiction arises because the analogy is not exact: reflection at a rigid acoustic boundary causes a phase shift of π, and a pressure release boundary causes zero phase lag. In contrast, due to evanescent near fields, reflection from the free end of a beam causes a phase lag of $\pi/2$. A more closely analogous acoustic system is therefore a dipole close to a boundary with a pure imaginary reflection coefficient giving 90° phase lag. It has been shown by numerical simulation (not shown here) that the power output in such a system is almost independent of the angle. Thus, on closer examination this result does not contradict the analogy.

VII. CONCLUDING REMARKS

Exact, closed-form expressions for the point force and moment mobility have been provided. The edge-excited plate can therefore be added to the relatively small list of plate structures for which exact analytical solutions are available-.Such solutions have application to testing approximate methods for example. It has been proved that the point mobility at a free edge is pure real like that of an infinite plate. The cross mobility cannot be solved in the same way, although the conjecture of Kauffmann's that the real and imaginary parts are equal has been proved. Therefore, it can be stated that the phase of the cross mobility is always $\pi/4$. The remaining numerical integration required to obtain its magnitude is significantly simplified. A further result, which is not intuitively obvious, is that real parts of the tangential and normal moment mobilities are the same, and that a moment with its axis in the plane of the plate will input the same power irrespective of its orientation.

ACKNOWLEDGMENTS

This work was partially funded by the University of Liverpool, and EC Framework V sponsored Project No. GRD1-1999-10785.

[1] E. Eichler, "Plate-edge admittances," J. Acoust. Soc. Am. **36**, 344–348 (1964).

[2] M. I. Mogilevski, "Point impedances of rectangular plates and vibration isolation of machines," J. Sound Vib. **161**(2), 213–225 (1993).

[3] R. Gunda, S. M. Vijayakar, R. Singh, and J. E. Farstad, "Harmonic Green's functions of a semi-infinite plate with clamped or free edges," J. Acoust. Soc. Am. **103**, 888–899 (1998).

[4] C. Kauffmman, "Input mobilities and power flows for edge-excited, semi-infinite, plates," J. Acoust. Soc. Am. **103**, 1874–1884 (1998).

[5] L. V. Ahlfors, *Complex Analysis* (McGraw–Hill, New York, 1979), Chap. 3, pp. 70–71.

[6] L. Cremer, M. Heckl, and E. E. Ungar, *Structureborne Sound* (Springer, Berlin, 1973), Table IV.1.

Interactions between a spherical elastic shell and acoustic waves from a water-entry moving source

M. Lee[a)]
Department of Mechanical Engineering, Sejong University, Seoul 143-747, Korea

(Received 28 June 2002; revised 20 December 2002; accepted 24 January 2004)

A possible interaction between the acoustic waves, which are generated from a water-entry body (moving source), and a submerged elastic shell is investigated theoretically within the scope of linear theory. The incident wave is defined from the ballistic wave model. The transient interaction is solved through extension of a method formulated for the excitation from a stationary source in an infinite domain. Numerical examples for the incident wave forms and corresponding shell responses are given to illustrate the effect of a moving source on the structure response. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1701900]

PACS numbers: 43.40.Fz [ANN] Pages: 2083–2090

I. INTRODUCTION

When a solid body penetrates water, two kinds of pressure waves are observed. In the early stage, *ballistic waves* are produced through the kinetic energy transfer of the body into water.[1–3] In the late stage of the penetration, the open cavity is necked down and becomes a bubble.[4] Then it oscillates to radiate the *collapse waves*.[5] It is of interest to note that the early stage ballistic waves come from a *moving source*, while the collapse waves come from a *stationary point source*. Hence the collapse waves are similar to what is seen in the underwater explosions. The point source wave due to underwater explosion has been extensively examined from the late 1940s.[6]

The generation of ballistic waves by a water-entry solid body was investigated more than 60 years ago by Mcmillen,[1,2] who used a spark shadow-graph of shock waves for entry of spheres into water. Leslie[5] measured the underwater sound produced by firing bullets into water, and was interested in the ballistic waves radiated by a bullet as well as the collapse waves created by the closed cavity (bubble) oscillations in the late stage of the penetration. Theoretical studies for calculating ballistic waves begin in the late 1960s with the work of Bristow and Lundenberg.[7] Yurkovich[8] used potential theory to describe the fluid motion, by placing a moving point source along the trajectory in the fluid.

These waves generated propagate underwater and eventually interact with submerged structures. In general, the transient interaction of acoustic shock waves generated from a *stationary* point source with a structure submerged in an infinite domain has been a topic under intensive study because of its important application area, e.g., naval vessels against underwater explosion and cavitation.[9–12] In the 1990s, these analytical solutions were used by researchers for validation of computer models involving fluid-structure interactions.[13] It can be noted that in the evaluation of the transient interactions between the incident waves from a stationary source and a submerged shell, the incident waves are provided from empirical equations.[6]

Although the wave generation process from a water-entry body and following wave interaction with a submerged structure was not studied much compared to the stationary point source, it has received some attention recently. Until now, however, no comprehensive analytical or numerical studies have been reported on the transient interactions between the water-entry ballistic waves and submerged structures. One of the important characteristics of these incident waves is that they are actually originated from distributed point sources (a *moving* source) along a trajectory with suitable time delay, corresponding to the motion of the source. In a previous work,[3] we have presented some theoretical and numerical results on ballistic waves generated by a solid body penetrating into water. Hence a goal of this paper is primarily devoted to the development of a method of calculating the solution of a transient interaction between the incident ballistic waves and a submerged elastic shell. For the incident wave field, we restrict our attention to the specific characterization of the pressure signatures at selected locations in the fluid. The transient response of a shell is solved through extension of a method previously formulated for the excitation from a *stationary* source, which employs the classical separation of variables, series solutions, and Laplace transform techniques.[9]

II. GENERATION OF INCIDENT WAVES

Consider the ballistic source-structure geometry as shown in Fig. 1. A solid body penetrates water in the $+z$ direction at an entry speed of v_i. During the penetration, pressure waves are generated through kinetic energy transfer. The ballistic waves originated from the moving solid eventually interact with a submerged elastic shell. Before proceeding with a description of the wave–structure interaction, a brief discussion of the generation of the incident wave will be necessary here. In order to have consistency throughout this paper, parameters are dimensionless,

a)Electronic mail: mlee@sejong.ac.kr

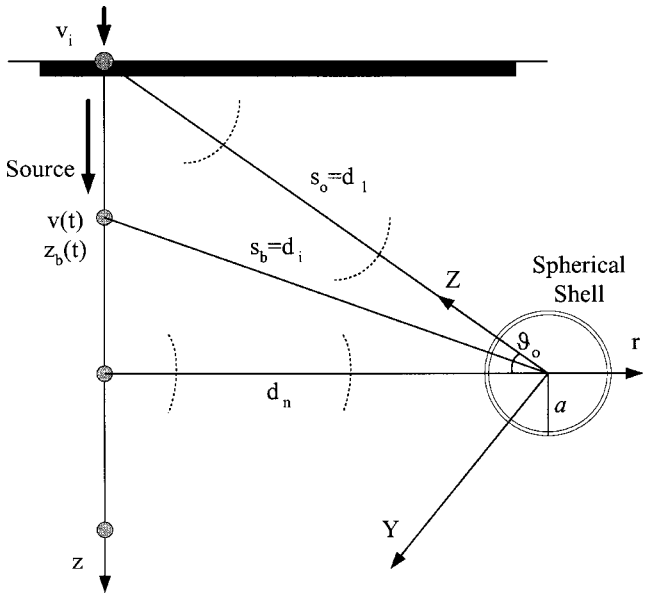


FIG. 1. Geometry of a submerged shell and ballistic waves from a moving body.

$$V=\frac{v_p}{c},\quad T=\frac{ct}{a},\quad Z=\frac{z}{a},\quad \Phi=\frac{\varphi}{ca},$$
$$\Pi=\frac{p}{\rho c^2},\quad \beta=\frac{\rho A_0 a C_d}{2m}, \tag{1}$$

where v_p is the moving source velocity, c the sound speed of water, t the time, a the radius of a submerged shell structure, z the penetration depth, φ a potential function, p the fluid pressure, ρ the water density, A_0 the projected area of the water-entry solid, C_d the drag coefficient, and m the mass of the water-entry solid. Here time is normalized, such that the incident wave completely engulfs the submerged shell at $T=2$. The dynamics of an entry body is then described using Newton's second law,

$$\frac{dV}{dT}=-\frac{\rho A_0 a C_d}{2m}V^2=-\beta V^2. \tag{2}$$

The rate of change of kinetic energy $((1/2)mV^2)$ of the entry solid with respect to depth is one of the important parameters since it is directly related to the magnitude of the incident ballistic waves. Using the velocity decay constant β, the rate of change of kinetic energy with respect to depth is given by

$$\frac{dE}{dZ_b}=\frac{mc^2}{a}\beta V^2, \tag{3}$$

where Z_b and β are the dimensionless penetration depth of a water-entry body and the velocity decay constant, respectively.

The pressure field radiated by a ballistic source (BS) is obtained by using the analysis of a moving point source. Adopting the small disturbance approximation and inviscid nonconducting flow, the fluid can be described in terms of a potential function, Φ, which satisfies the wave equation,[14]

$$\nabla^2\Phi=\frac{\partial^2\Phi}{\partial T^2}. \tag{4}$$

The effect of a moving solid is approximated by distributed point sources along a trajectory as shown in Fig. 1. For an arbitrary source velocity, the potential is the total radiation summed over a line source distribution along the z axis. The pressure generated by a BS is then obtained using Bernoulli's equation. Although the theory is developed to take the dynamic pressure term in the unsteady Bernoulli's equation into consideration, it is generally negligible at a far distant target such as

$$\Pi=-\frac{\partial\Phi}{\partial T}. \tag{5}$$

To satisfy the pressure-release boundary condition at the free surface, a mirror image is placed above the fluid with 180° out of phase. Hence, incident wave forms are obtained with a consideration of source and mirror images. If the entry-velocity is subsonic, it is not necessary to consider the time-inversion effect.[3,15] The time-inversion effect simply means that a wavelet emitted at early time arrives to the observer later than the second wavelet. This could occur when the source is moving toward the receiver at supersonic velocity. Now the time derivative of the potential becomes[14]

$$\frac{\partial\Phi}{\partial T}=-\frac{1}{2}\frac{BA_b}{acS_b}\frac{V}{1-V\dfrac{Z-Z_b}{S_b}}+\frac{1}{2}B^2\ln\left[\frac{Z+S_0}{Z-Z_b+S_b}\right], \tag{6}$$

where

$$A^2=\frac{dE_p/dZ_b}{\pi a\rho c^2\Pi_c},\quad B^2=\frac{\Pi_c}{c^2N}. \tag{7}$$

Here, S_0 denotes the initial distance between the water-entry point and measurement location in the fluid, while S_b is the distance estimated from the time-dependent source locations. A_b is the value of A estimated at Z_b, $\Pi_c=(p_0-p_c)/\rho c^2$, p_0 and p_c are the ambient and cavity inside pressure, respectively. N is a geometric dimensionless parameter that accounts for the range of influence of the divergence of the flow. To make the paper more complete, the derivation of the potential function and its derivative is provided in the Appendix. It can be noted that the time-domain approach used to predict the pressure wave form is perfectly suitable for a numerical computation. The procedure to determine the ballistic wave form at specific locations as a function of time is fully provided in the previous paper.[3] Having defined the dynamics of a solid and the associated ballistic waves, we proceed then to the transient interaction analysis between these incident waves and a submerged shell.

III. WAVE–STRUCTURE INTERACTIONS

Since the ballistic waves radiated by a moving source are continuous and originated from different locations along a trajectory, they (source and mirror images) have different contact points on the submerged spherical shell as shown in Fig. 1. In order to manage this difficulty in calculating the total response of a shell subjected to a distributed wave loading, the basic idea is that the incident waves are discretized into a series of elementary step pulses that originate from

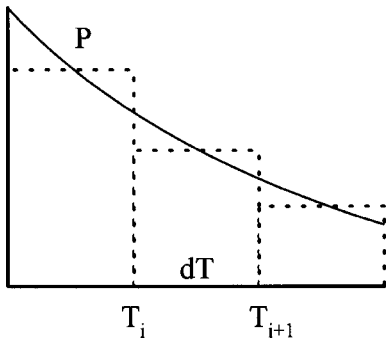


FIG. 2. The continuous ballistic waves (solid line) with a series of elementary pulses (broken line).

different source and mirror image centers. That is, the wave is replaced by a piecewise constant function as shown in Fig. 2. Then, the interaction of each elementary pulse with a shell is analyzed. To do this, a submerged elastic shell is oriented in such a way that each elementary pulse always strikes the vertex of the shell in the XYZ coordinate system as shown in Fig. 3. It is then necessary to apply a coordinate transformation technique to each response since the response of a shell at a specific location is to be estimated in the global xyz coordinate system.

Finally, the total response of the shell is obtained by adding the elementary response to each pulse emitted by each part of the source and mirror images with a suitable time delay corresponding to the motion of the solid. The detailed analysis for predicting the shell response subjected to a water-entry moving source will be described in the following sections of this paper.

A. Interaction with elementary pulse from a stationary source

First, we consider the interaction between an elastic shell and an elementary unit step pulse from a stationary source. Incident pressure waves of a unit step pulse are impinging on a spherical elastic shell of radius a as shown in Fig. 3. The XYZ system is located on the unperturbed center of the shell. Although it may be more accurate to use a spherical wave form as an incident wave, the plane wave form is used in this analysis. This is mainly because the error can be reduced by increasing the number of incoming elementary wave pulses (the number of moving point sources), and partially because for $D \geqslant 10$ the response is about the same as that for the case of a plane incident wave,[10] where D is the dimensionless distance between the source and the center of a shell. With no further difficulty, the incidence can be a spherical pulse. The surrounding fluid is taken to be isotropic and homogeneous. The structure is located at a far distance such that the incident wave is sufficiently weak and the linear acoustic wave theory is sufficient to describe the propagation of acoustic shock waves.

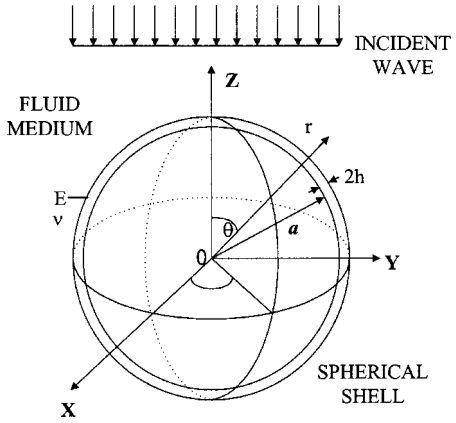


FIG. 3. Geometry of a spherical shell and an elementary incident pulse.

The shell is considered to be made of an isotropic elastic material. The properties of the shell are radius a, thickness $2h$, Young's modulus E, Poisson's ratio ν, and mass density ρ_s. The middle surface deflections of the shell are denoted by δ_r and δ_θ in the r and θ directions, respectively. The dimensionless parameters used in the wave–structure interaction analysis are given here,

$$w=\frac{\delta_r}{a}, \quad u=\frac{\delta_\theta}{a}, \quad R=\frac{r}{a}, \quad D=\frac{d}{a},$$
$$C^2=\frac{E}{[\rho_s(1-\nu)c^2]}, \quad M=\frac{\rho a}{2h\rho_s}. \tag{8}$$

The time scale is the same as in section III. M is equal to one-third of the ratio of the mass of the fluid with the same volume of the sphere to the mass of the shell. If the shell is in a neutral buoyancy state, $M=3$. The elastic and rigid-body motions of a shell are initiated by the incident waves that are then scattered by the shell. The shell also radiates waves into the fluid medium. So, the resultant pressure (Π) field in the surrounding medium is the sum of two parts: the incident pressure (Π^{inc}) and the scattered and radiated pressure (Π^{sra}) due to the motion of the shell.

Since the wave–structure interactions involve two mediums, two different governing equations, the Lagrange equation of shell and wave equation, must be solved simultaneously. From the physical background, the deflections of the shell and the wave pressure are continuous; thus they can be expanded in terms of Legendre functions,[15]

$$w(\theta,t)=\sum_{m=0}^{\infty} w_m(T)P_m(\cos\theta), \tag{9}$$

$$u(\theta,t)=\sum_{m=1}^{\infty} u_m(T)\frac{d}{d\theta}P_m(\cos\theta), \tag{10}$$

$$\Pi(R,\theta,t)=\sum_{m=0}^{\infty} \Pi_m(R,T)P_m(\cos\theta). \tag{11}$$

The Lagrange equations of the system are

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{w}}-\frac{\partial K}{\partial w}+\frac{\partial V}{\partial u}=f_i, \tag{12}$$

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{u}}-\frac{\partial K}{\partial u}+\frac{\partial V}{\partial w}=0, \tag{13}$$

where K and V are the kinetic and potential energies of the shell, respectively. The quantities required for these equations will be evaluated and expressed in terms of the gener-

alized velocities after substitution of the series.

Now the generalized force in Eq. (12) is given by the variation of the work done on the shell by the wave pressure, which will be provided by the analysis in the previous section. The scattered pressure, typically denoting the sum of scattered and radiated terms, should satisfy the dimensionless wave equation in spherical coordinates,

$$\frac{1}{R}\frac{\partial}{\partial R}\left(R^2\frac{\partial \Pi^{\mathrm{sra}}}{\partial R}\right)+\frac{1}{R^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial \Pi^{\mathrm{sra}}}{\partial\theta}\right)=\frac{\partial^2\Pi^{\mathrm{sra}}}{\partial T^2}. \tag{14}$$

The scattered pressure for the spherical wave equation for an outgoing wave is of the form of products of Legendre functions and modified spherical Bessel functions of the third kind. Using the Laplace transform and the wave equation for Π^{sra} and the shell equation of motion for the deflections allows us to determine the modal radial deflection of the shell. The description of the interaction analysis will be completed by including an important boundary condition, which is the continuity of the normal velocity at the structure–fluid interface,

$$\frac{\partial \Pi}{\partial R}=-\frac{\partial^2 w}{\partial T^2} \quad \text{at } R=1. \tag{15}$$

The wave pressure field can be written as the sum of two separable parts, i.e.,

$$\Pi(R,\theta,T)=\Pi^{\mathrm{inc}}(R,\theta,T)+\Pi^{\mathrm{src}}(R,\theta,T). \tag{16}$$

An arbitrary incident plane wave can be expressed by the following function:

$$\Pi^{\mathrm{inc}}(Z,T)=\Pi^{\mathrm{inc}}(R,\theta,T)=F(Z-1+T)H(Z-1+T)$$
$$=F(R\cos\theta-1+T)H(R\cos\theta-1+T), \tag{17}$$

where

$$H(R\cos\theta-1+T)=\begin{cases}0 & \text{if } R\cos\theta-1+T<0\\ 1/2 & \text{if } R\cos\theta-1+T=0\\ 1 & \text{if } R\cos\theta-1+T>0\end{cases} \tag{18}$$

Once an arbitrary function, F, of its argument is defined from the ballistic wave generation model, the wave–structural interaction solutions due to the elementary wave pulse can be obtained by using the same procedure as in the previous paper.[9]

In Fig. 4, it is demonstrated first that the present program can accurately reproduce the Huang's results. Since the most significant phenomena occurs in the early time response, results up to nondimensional time 6 are compared here.

B. Interactions with piecewise continuous waves from a moving source

It should be understood that, in the case of a moving source, the coordinates of the moving point source and mirror images are time-dependent, which induces a variation of the angle of θ. Hence, each elementary pulse interacts with the shell at different locations as shown in Fig. 5. To eliminate the complexity of the analysis, a coordinate transformation technique is adopted such that each elementary pulse always strikes the vertex $(Z=1)$ of the shell in the XYZ system. In order to make this more concrete, the procedures are as follows. Given the source for the line-source representation of a particular bullet entry, one divides the line into fixed segments of length and calculates the acoustic field Φ generated by each segment considered individually. Then one orients an ensemble of spherical shells of common radius and centered at a common point such that the axis of each shell passes through the center of each different segment, thereby defining an ensemble of axis angles, θ_i. Next, one uses Huang's single-source solution (section III.A) to determine, for the ith segment, the radial shell velocity $v_i(\vartheta_i,t)$ at a meridional angle ϑ_i selected such that $\theta_i+\vartheta_i=\Theta$, where Θ is a prescribed angle. Then this process is repeated (with index j) for the set of mirror-image negative-source line segments that account for the presence of the free surface. Finally, one obtains the radial shell velocity at the prescribed angle Θ by superposition as $v(\Theta_i,t)=\Sigma\, v_i(\vartheta_i,t)+\Sigma\, v_j(\vartheta_j,t)$.

Next it is necessary to describe the arrival time of each elementary pulse emitted from the specific locations of each

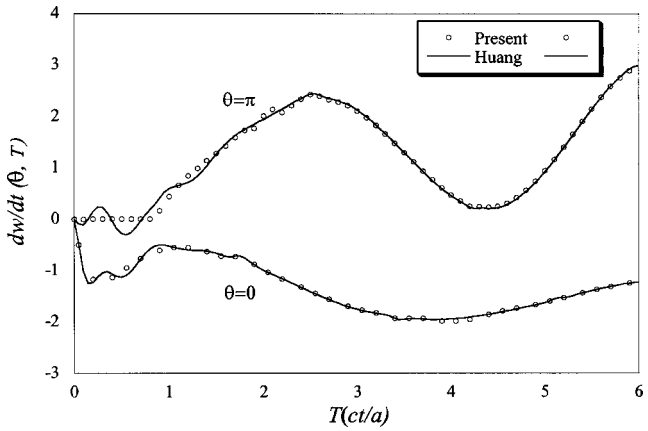


FIG. 4. Comparison of shell radial velocities between present and Huang's results.

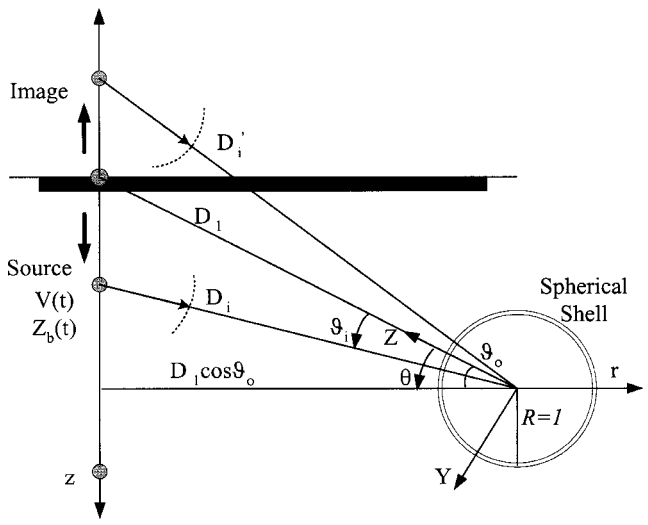


FIG. 5. Geometry of the source mirror images.

source and images to the structure. This is estimated by taking into account the time delay corresponding to the motion of a source and the speed of sound. We are reminded that the origin of the time coordinate for each interaction is adjusted such that the first elementary step pulse strikes the vertex of the shell ($R=1$, $\vartheta=\vartheta_o$) at $T=0$. Noting that $V=dZ/dT$ and assuming the propagation speed of each pulse as the acoustic speed c, the arrival time difference between the first pulse and each elementary pulse is

$$\Delta T=\left(\int_0^{Z_b}\frac{1}{V(Z)}dZ+D_i\right)-D_1, \tag{19}$$

where V and Z_b are the time-dependent source velocity and moving distance, as described in the section II. The distance between the shell center and the ith source is D_i. Equation (19) is enough to determine the arrival time of each pulse. For a mirror image, the traveling distance of each emitted wave is estimated from the geometry in Fig. 1,

$$D_i=\sqrt{(D_1\cos\vartheta_0)^2+\left(\int_0^{Z_b}\frac{1}{V(Z)}dZ+D_1\sin\vartheta_0\right)^2}. \tag{20}$$

The source and image are taken into account separately since their first contact location on the submerged shell is not the same as time goes by.

The pressure-release boundary condition is satisfied at the free surface. It is clear that the free surface reflection consists of two parts: the reflection of the incident wave and that of the pressure field scattered and radiated by the shell. The former has already been included as an incident wave using the image source method. If the latter is considered, the problem becomes that of multiple scattering by two shells. However, this is, in general, negligible because the scattered and radiated pressure waves emanating from the shell are small for a case in which the shell is several diameters away from the free surface. As far as the structural response is concerned, the problem can be reduced to that of a structure immersed in an infinite fluid and impinged upon two successive incident waves.

C. Solution procedures

The objective of this section is to outline the procedure employed in calculating the transient response of a submerged elastic shell subjected to ballistic waves that are radiated by a moving source. The three basic steps are as follows: First, the motion of a moving source and image is decomposed in space into constant distances ΔZ, along a trajectory, where the dynamics of the source are determined using the momentum equation. Second, with determination of the arrival time and incident angle of each elementary step pulse, the corresponding incident ballistic wave forms are constructed for the source and mirror images individually. Finally, the transient interactions analysis between each elementary pulse and an elastic shell are carried out, and the coordinate transformation technique is applied to obtain the total response of the shell due to the distributed incident wave loadings from a moving source.

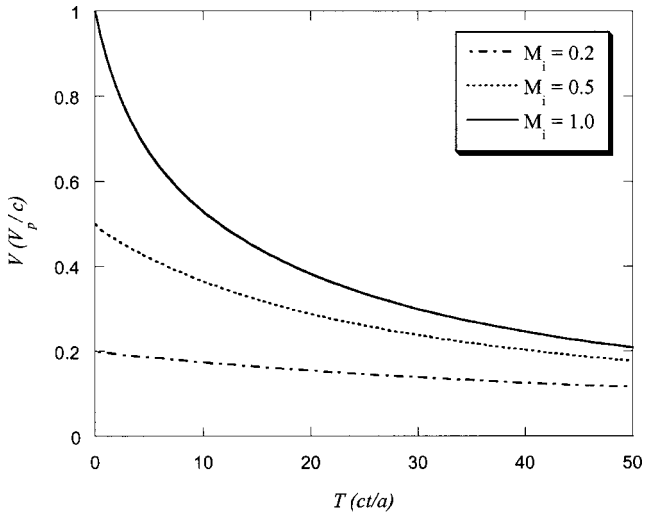


FIG. 6. Time histories of penetration velocity for the entry of a 2 cm sphere. The time scale was normalized such that the incident wave propagates one radius of a shell at $T=1$, $\Gamma=M_iD_1\cos\vartheta_o$, $D_1\cos\vartheta_o=6$, $\vartheta_o=45°$.

IV. RESULTS

A. Incident ballistic waves

Numerical calculations are carried out for the case of a sphere of 2-cm-diam penetrating water. Figures 6 and 7 show the predicted penetration velocity and depth as the dimensionless parameter, $\Gamma=M_iD_1\cos\vartheta_o$, varies from 1.2 to 6, with $D_1\cos\vartheta_o=6$ and $\vartheta_o=45°$. The entry Mach number of the source, M_i, is varied from 0.2 to 1.0. As shown in Fig. 6, the deceleration of the penetration velocity after an entry is more significant for $\Gamma=6$ and less significant for $\Gamma=1.2$. For the case of $\Gamma=1.2$, the source has lost only 10% of its initial velocity up to $T=10$. Obviously, the dimensionless penetration depth data are important since they govern the loading duration and magnitudes of the distributed pulses through the time-dependent relative coordinates between the source and target structure. This is because the origin of the time coor-

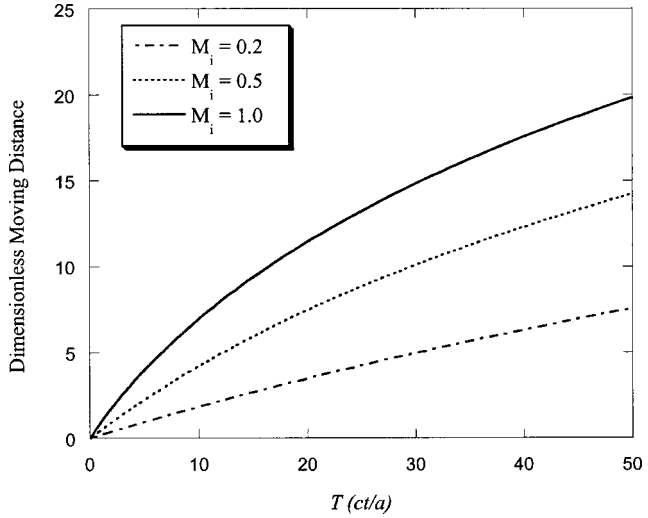


FIG. 7. Time histories of penetration depth for the entry of a 2 cm sphere, $\Gamma=M_iD_1\cos\vartheta_o$, $D_1\cos\vartheta_o=6$, $\vartheta_o=45°$.

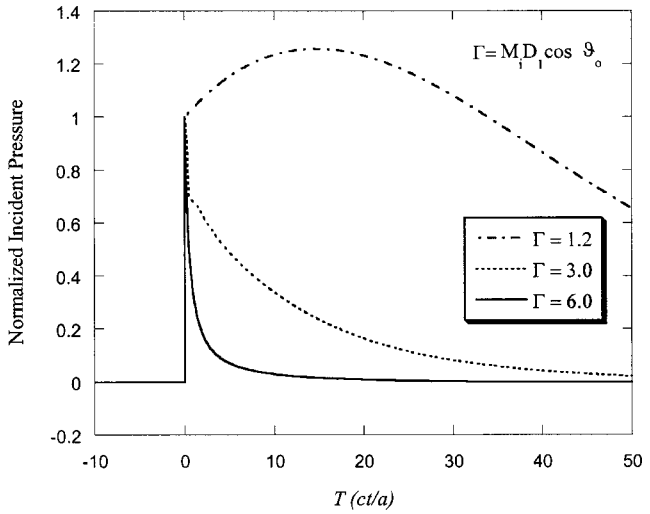


FIG. 8. Time history of incident ballistic waves for the entry of a 2 cm sphere; $\Gamma = M_i D_1 \cos \vartheta_o$, $D_1 \cos \vartheta_o = 6$, $\vartheta_o = 45°$. For the cases of $\Gamma = 3$ and 6, after an initial peak, a roughly exponential decay in the amplitude is observed. For the case of $\Gamma = 1.2$, the amplitude remains constant (≈ 1) until approximately $T = 40$. A calculation of the wave form allows us to determine whether a submerged shell is subjected to a step wave or an exponential decay wave.

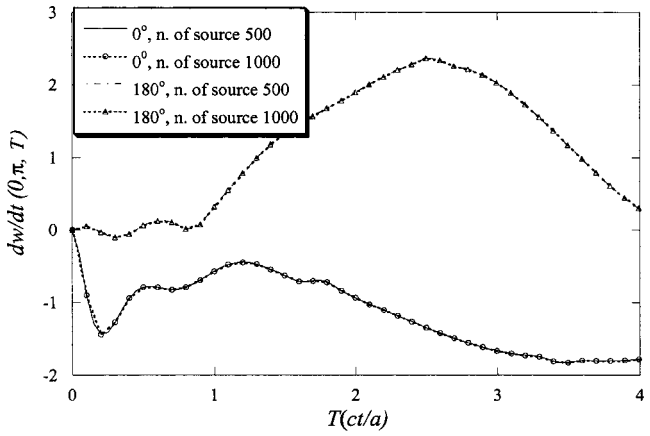


FIG. 9. Radial velocities of the shell for different number of sources.

dinate is again adjusted and normalized such that the incident wave first impinges upon the shell at $\vartheta = \vartheta_o$ and $T = 0$, and it propagates one radius into the shell at $T = 1$.

A plot of a typical wave form radiated by a sphere of a 2 cm diameter, as the dimensionless parameter Γ varies from 1.2 to 6, is displayed in Fig. 8. The signatures are estimated at a depth and width each of twelve radii (refer to Fig. 1). As seen in the figure, the incident wave is normalized by the initial peak pressure magnitude. The range of value Γ was chosen in a way to illustrate different incident wave characteristics. For the cases of $\Gamma = 3$ and 6, a decay wave is observed. For the case of $\Gamma = 1.2$, however, the amplitude remains constant (≈ 1) until $T = 40$. Note that the incident wave completely engulfs the shell at $T = 2$. In this case, a uniform incident pulse, instead of a decay pulse, is supposed to impinge the submerged shell. It can be emphasized that the characteristics of the incident waves can be successfully examined with a variation of Γ. In fact, the distinguishable wave form for $\Gamma = 1.2$ is due to the fact that this source decelerates slowly while the distance between the source and receiver becomes shorter, which radiates a steady pulse.[3] In conclusion, a calculation of the wave form allows us to determine whether a submerged shell is subjected to a uniform wave or a decay wave. It can also be emphasized that the characteristics of the incident wave are determined in terms of the size of the target structure. That is, it is important to estimate the amplitude of the incident wave as it engulfs the shell and propagates further away.

B. Transient responses

The transient responses of a steel spherical shell subjected to the incident waves defined in the previous section are analyzed in this section. The thickness-to-diameter ratio h/a is taken to be 0.01. The properties of the steel shell are $E = 110\,\text{GPa}$, $\upsilon = 0.3$, and $\rho_S = 7780\,\text{kg/m}^3$. The acoustic wave speed in water is 1480 m/s. Even though the shell is not in the neutral buoyancy state for these material properties, the problem is assumed to be linear and thus the effect of the buoyancy force is decoupled from the wave–shell interaction.[9]

It can also be noted that a sensitivity study was carried out in this study to determine the sufficient number of point sources. Figure 9 shows the early time responses of the shell radial velocity at $\Theta = 0$ and π for $\Gamma = 3.0$ with the number of sources 500 and 1000. As shown in Fig. 9, the early time responses, which seem to be most important, are almost the same between the solutions. From this, the number of sources is fixed at 500. As can be expected, the number of sources depends on the function of incident wave. If the incident wave is close to the uniform wave, the number of sources may be reduced significantly without any difference in the response.

Figure 10 shows the time histories of the shell radial velocity at $\Theta = 0$ and π due to the previously described incident waves. For comparison purposes, the step incident result from a stationary source ($\Gamma = 0$, Huang's theoretical prediction[9]), which is assumed to be located at the water-entry distance in an infinite medium, is also displayed. The magnitude of the step pulse is one, as obtained by Huang. For the stationary source case, the radial velocities at $\Theta = 0$ and π oscillate about a rigid body velocity at a large value of T. For the case of ballistic incidence from a moving source, however, the velocity at $\Theta = 0$ initially approaches that of a rigid body velocity and subsequently goes to zero. That is, the velocity is a damped oscillatory function of T at larger values of T. It can be seen that the peak elastic velocity for $\Gamma = 3.0$ drops down rapidly in a very short duration of time. As expected from physical considerations, this is contributed from the distributed incidence angle resulting from the high moving speed of a source. That is, the incident angle varies in proportion to the increased speed of the source. This is the general characteristics of the wave–structure interactions between a submerged target structure and a moving source. The same trend is also observed for the velocity response of the shell at $\Theta = \pi$.

The time histories of the shell radial velocity at $\Theta = \pi/2$ and $3\pi/2$ due to the previously described incident waves are shown in Fig. 11. For the case of a plane step incident wave

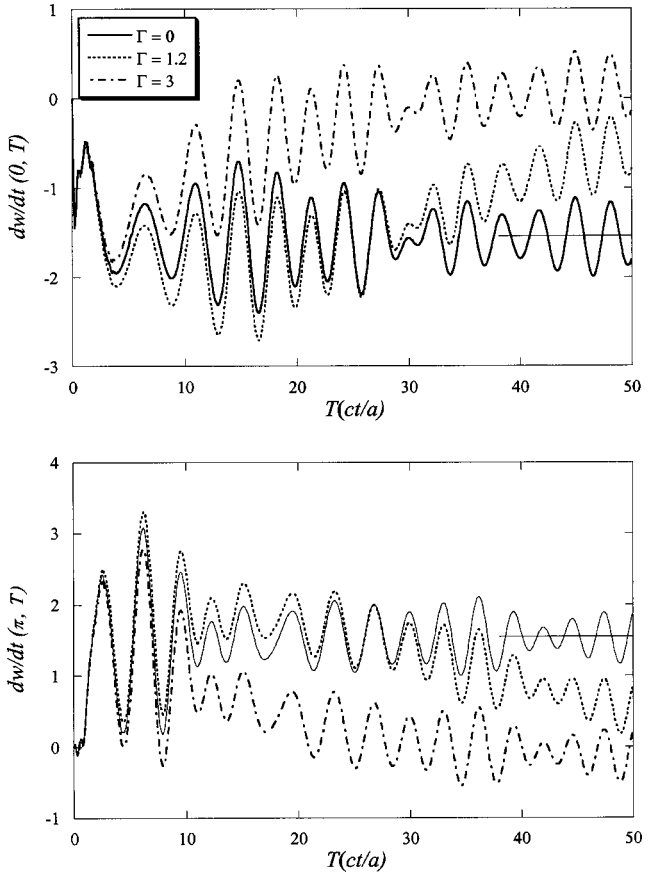


FIG. 10. Radial velocity of the shell estimated at $\Theta=0$ and π for the entry of a 2 cm sphere; $\Gamma=M_iD_1\cos\vartheta_o$, $D_1\cos\vartheta_o=6$, $\vartheta_o=45°$. The plane incident wave case from a stationary solution in an infinite domain ($\Gamma=0$, Huang's theoretical prediction), where for large values of T velocities oscillate about the value of rigid-body, is compared with the moving source solutions.

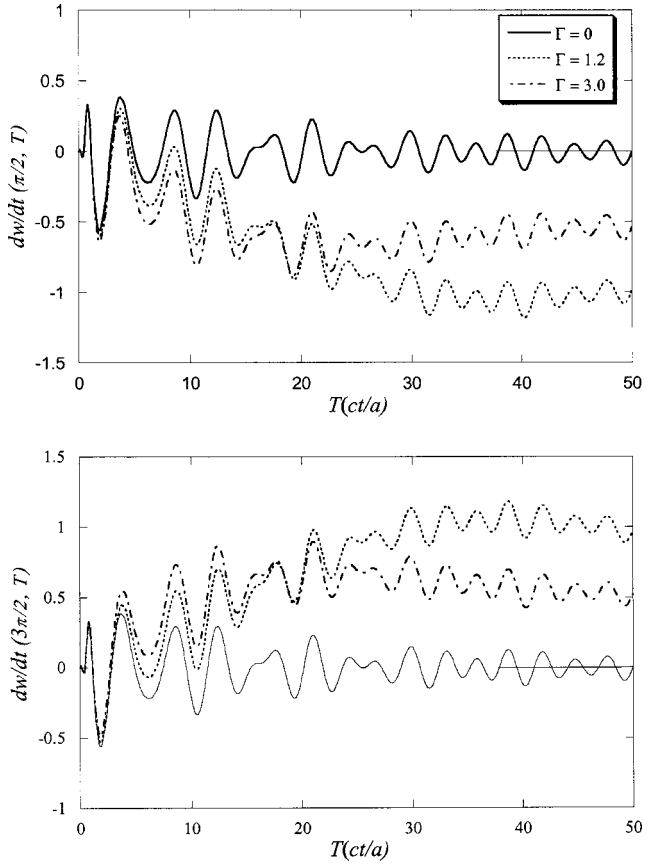


FIG. 11. Radial velocity of the shell estimated at $\Theta=\pi/2$ and $3\pi/2$ for the entry of a 2 cm sphere; $\Gamma=M_iD_1\cos\vartheta_o$, $D_1\cos\vartheta_o=6$, $\vartheta_o=45°$. The plane incident wave case from a stationary solution in an infinite domain ($\Gamma=0$, Huang's theoretical prediction), where for large values of T velocities oscillate about the value of zero, is compared with the moving source solutions.

from a stationary source ($\Gamma=0$, Huang's theoretical prediction[9]), the radial velocity at $\Theta=\pi/2$ oscillates with diminishing amplitude about zero. For the case of ballistic incidence, however, the velocities increase and reach their asymptotic values at about $T=20$ and 30 for $\Gamma=1.2$ and $\Gamma=3.0$, respectively. They will vanish if further calculations ($T>60$) are carried out, as the magnitude of incident wave decays. As long as the initial stage response is concerned (actually most important), it is not necessary to calculate longer period of time. It is interesting to note that a time-dependent unsymmetrical loading from a moving source causes different shell responses between at $\Theta=\pi/2$ and $3\pi/2$, while the same responses are observed in the case of a stationary source. Consequently, these responses of the shell are the characteristics of interaction between a moving source and a submerged shell. It can be addressed that two major parameters governing the interaction between a moving source and a submerged shell are a time-dependent contact angle, which is related to the source velocity, and an incident wave function (for example, decay constant).

V. CONCLUSIONS

In this paper, a complete methodology for the transient interaction between a submerged elastic shell and ballistic waves from a water-entry solid has been presented. The incident wave is defined from the ballistic wave model, and its interaction with a submerged structure is solved through extension of a method previously formulated for the stationary point source excitation in an infinite domain. Numerical examples have been presented to illustrate the effect of the type of a moving source on the structural response. Since one of the important characteristics associated with a moving source is that it involves a distributed wave loading to a submerged structure, the structural response to a moving source is different from that to a stationary source. The major parameters governing the extent of the interactions are the time-dependent contact angle and the incident wave form. The present study is applicable to sonor systems. Since it is a ballistic (moving) source, the acoustic field measured in the source system is unique.

ACKNOWLEDGMENT

This work was supported by Korea Research Foundation Grant No. KRF-2001-2001-041-E00019.

APPENDIX: DERIVATIVES OF POTENTIAL

The derivation of the velocity potential is provided in the Appendix in a dimensional form.[14] As described in this

paper, the effect of the bullet and cavity on the fluid is approximated by the action of a line of sources. Then, the resulting flow field is symmetric about the z axis, and the potential can be expressed as

$$\varphi(z,\varpi,t) = -\int_0^{z_b(\tau)} \frac{\xi(\xi,\tau)}{r}\, d\xi, \tag{A1}$$

where ξ is the distance along the trajectory, r the distance between point ξ and (z,ϖ), z_b the bullet position, ξ the source strength at ξ. For the finite sound speed, the integral must be evaluated along the line $\tau + r/c = \text{constant}$, where τ is the retarted time given by $\tau = t - r/c$.

The source strength is estimated by a method based on the conservation of energy. That is, the rate of loss of kinetic energy of a source is converted into the kinetic energy and potential energy in the fluid section, which results in

$$\varphi(z,\varpi,t) = -\frac{1}{2}B\int_0^{z_b(\tau)} \frac{1}{r}\left\{ A(\xi) - B\left[t - \frac{r}{c} - t_b(\xi) \right] \right\} d\xi\varphi. \tag{A2}$$

The nondimensional form of A and B is given in Eq. (7) in this paper. Note that the retarded time τ was replaced by $t - r/c$.

To compute the pressure, the terms in Bernoulli's equation [Eq. (5)] must be evaluated. The unsteady term, $\partial\varphi/\partial t$, in the Bernoulli's equation can be expressed in a simple form using Leibnitz's rule for the differentiation of integrals. Then Eq. (A2) becomes

$$\frac{\partial\varphi}{\partial t} = -\frac{1}{2}\left[\frac{BA_b}{s_b}\frac{\partial z_b(\tau)}{\partial t} \right] + \frac{1}{2}B^2\int_0^{z_b(\tau)} \frac{1}{r}\, d\xi, \tag{A3}$$

where s_b is the distance between the bullet and the point (z,ϖ) and A_b the value of A evaluated at z_b. The chain rule for the differentiation gives

$$\frac{\partial z_b(\tau)}{\partial t} = \frac{\partial z_b(\tau)}{\partial \tau}\frac{\partial \tau}{\partial t} = v\frac{\partial \tau}{\partial t}, \tag{A4}$$

where v is the source velocity. From the following relation,

$$\frac{\partial \tau}{\partial t} = 1 - \frac{1}{c}\frac{\partial s_b(\tau)}{\partial t} = 1 + \frac{1}{c}\frac{z - z_b}{s_b}\frac{\partial z_b(\tau)}{\partial t}, \tag{A5}$$

Eq. (A4) becomes

$$\frac{\partial z_b(\tau)}{\partial t} = v\left[1 + \frac{1}{c}\frac{z - z_b}{s_b}\frac{\partial z_b(\tau)}{\partial t} \right]. \tag{A6}$$

Solving for $\partial z_b(\tau)/\partial t$ gives

$$\frac{\partial z_b(\tau)}{\partial t} = \frac{v}{1 - \frac{v}{c}\frac{z - z_b}{s_b}}. \tag{A7}$$

With Eq. (A7) and solving the integral, Eq. (A3) becomes

$$\frac{\partial\varphi}{\partial t} = -\frac{1}{2}\frac{BA_b}{s_b}\frac{v}{1 - \frac{v}{c}\frac{z - z_b}{s_b}} + \frac{1}{2}B^2\ln\left[\frac{z + s_0}{z - z_b + s_b} \right], \tag{A8}$$

where s_0 is the distance between the entry point and the point (z,ϖ). The nondimensional form of Eq. (A8) is provided in Eq. (6).[16]

[1] J. H. Mcmillen, "Shock wave pressures in water produced by impact of small spheres," Phys. Rev. **68**, 198–209 (1945).
[2] J. H. Mcmillen and E. N. Harvey, "A spark shadow graphic study of body waves in water," J. Appl. Phys. **17**, 541–555 (1946).
[3] M. Lee, R. G. Longoria, and D. E. Wilson, "Ballistic waves in high-speed water entry," J. Fluids Struct. **11**, 819–844 (1997).
[4] M. Lee, R. G. Longoria, and D. E. Wilson, "Cavity dynamics in high-speed water entry," Phys. Fluids **11**, 540–550 (1997).
[5] C. B. Leslie, "Underwater noise produced by bullet entry," J. Acoust. Soc. Am. **36**, 1138–1144 (1964).
[6] R. H. Cole, *Underwater Explosions* (Princeton University Press, Princeton, 1948).
[7] R. J. Bristow and J. F. Lundenberg, "Hydraulic Ram," FADS-66, Boeing Co., Seattle, 1969.
[8] R. Yurkovich, "Hydraulic Ram: A fuel tank vulnerability study," Report No. G964, McDonell Douglas Co., St. Louis, 1969.
[9] H. Huang, "Transient interaction of plane acoustic waves with a spherical elastic shell," J. Acoust. Soc. Am. **45**, 661–670 (1969).
[10] H. Huang, Y. P. Lu, and Y. F. Wang, "Transient interaction of spherical acoustic waves with a spherical elastic shell," J. Appl. Mech. **38**, 71–74 (1971).
[11] P. Zhang and T. L. Geers, "Excitation of a fluid-filled, submerged spherical shell by a transient acoustic wave," J. Acoust. Soc. Am. **93**, 696–705 (1993).
[12] M. A. Sorague and T. L. Geers, "Response of empty and fluid-filled, submerged spherical shells to plane and spherical, step-exponential acoustic waves," Shock Vib. **6**, 147–157 (1999).
[13] M. Lee, B. Lee, and S. Y. Lee, "On the analysis methods for interactions of a submerged elastic shell and acoustic shocks," in *Hydrodynamics IV: Theory and Applications*, edited by Y. Goda, M. Ikehata, and K. Suzuki (2000), pp. 441–446.
[14] E. A. Lundstrom, "Fluid dynamic analysis of hydraulic ram," NWC TP 5227, Naval Weapons Center, China Lake, CA, 1988.
[15] Y. H. Berthelot and I. J. Busch-Vishniac, "Thermoacoustic radiation of sound by a moving laser source," J. Acoust. Soc. Am. **81**, 317–327 (1987).
[16] P. Mann-Nachbar, "The interaction of an acoustic wave and an elastic spherical shell," Quart. J. Appl. Math. **15**, 83–93 (1957).

Analytical approach for sound attenuation in perforated dissipative silencers

A. Selamet,[a] M. B. Xu, and I.-J. Lee
Department of Mechanical Engineering and The Center for Automotive Research, The Ohio State University, Columbus, Ohio 43210-1107

N. T. Huff
Owens Corning, Novi, Michigan 48377

(Received 26 August 2003; revised 2 February 2004; accepted 10 February 2004)

A two-dimensional analytical solution is developed to determine the acoustic performance of a perforated single-pass, concentric cylindrical silencer filled with fibrous material. To account for the wave propagation through absorbing fiber and perforations, the complex characteristic impedance, wave number, and perforation impedance are employed. With expressions for the eigenvalues and eigenfunctions of sound propagation in the perforated dissipative chamber, the transmission loss is obtained by applying a pressure and velocity matching technique. The results from the analytical method are then compared with both experiments and numerical predictions based on the boundary element method (BEM), showing a reasonable agreement. The effects of geometry, fiber properties, and perforation porosity on the acoustic attenuation performance are discussed in detail. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1694994]

PACS numbers: 43.50.Gf, 43.20.Mv [ANN] Pages: 2091–2099

I. INTRODUCTION

Dissipative silencers are used to control noise in automotive exhaust systems due to their broadband attenuation characteristics (particularly at mid to high frequencies) and low back pressure. Thus far, a number of numerical techniques have been used to quantify their acoustic behavior, including the finite element[1,2] and boundary element methods.[3–6] Given the computational effort for sufficiently accurate numerical results, however, an analytical approach may be a viable alternative for silencers with simple cross sections, such as circular. In terms of bulk-reacting model, the acoustic attenuation in infinite rectangular and circular lined ducts was analyzed by Scott;[7] the characteristics of sound transmission in infinite rectangle, annular, and circular lined ducts with mean flow were investigated by Ko;[8] the effects of a perforated screen and mean flow on the modal attenuation rates in infinite circular lined ducts were examined by Nilsson and Brander,[9] and Cummings and Chang,[10] respectively. Cummings and Chang[11] further investigated a finite-length dissipative silencer including mean flow and obtained the transmission loss by a mode-matching technique using the continuity of acoustic pressure and axial particle velocity across the silencer discontinuities. Peat[12] obtained the transfer matrix for a dissipative expansion chamber by a ''low-frequency approximation,'' which is accurate at low frequencies. For a circular dissipative expansion chamber with the fibrous material separated from the central airway by a perforated screen, Kirby[13] obtained the transmission loss using a technique similar to the low-frequency approximation by including additional terms in the series expansions for the Bessel and Neumann functions. This approach, however, remains confined to relatively low frequencies due to the inherent limitation of expansions with a finite number of terms. Wang[14] developed a one-dimensional (1D) decoupling method to study the acoustic attenuation of a resonator with fiber in the expansion chamber. Selamet *et al.*[4] investigated the sound attenuation in single-pass concentric perforated dissipative silencers by a 1D analytical approach, BEM, and experiments. Munjal and Thawani[15] examined the effect of a perforated protective plate or a thin impervious layer on the acoustic performance of lined circular ducts and parallel-baffle mufflers. For both cases, they presented analytical models with bulk-reacting as well as locally reacting absorptive linings, while accounting for the effect of grazing flow on the impedance of perforated plate. Munjal[16] also investigated a pod silencer consisting of a lined circular duct with a cylindrical pod inside, and obtained the four-pole parameters and transmission loss. By a pressure and velocity matching technique originated in Ref. 17, a closed-form, two-dimensional analytical solution was developed to investigate the sound attenuation in a circular dissipative expansion chamber.[18]

In Ref. 18, no perforated screen was considered in the dissipative chamber, leaving fiber in direct contact with the gas in the central airway. However, to retain the fibrous material, a perforated screen is usually placed between the fiber and the central airway, which impacts the acoustic performance of the silencer. The acoustic impedance of a porous layer with perforated facing was investigated by Ingard and Bolt.[19] Kirby and Cummings[20] obtained the acoustic impedance of perforated plates by examining experimentally two types of perforations, circular and louvered plates, with and without porous backing. Empirical formulas and semiempirical predictions were provided for the case without, and with fiber backing layer, respectively. The objective of the present study is to extend an earlier work[18] by introducing the perforated screen into a single-pass, concentric, dissipative cy-

[a]Electronic mail: selamet.1@osu.edu

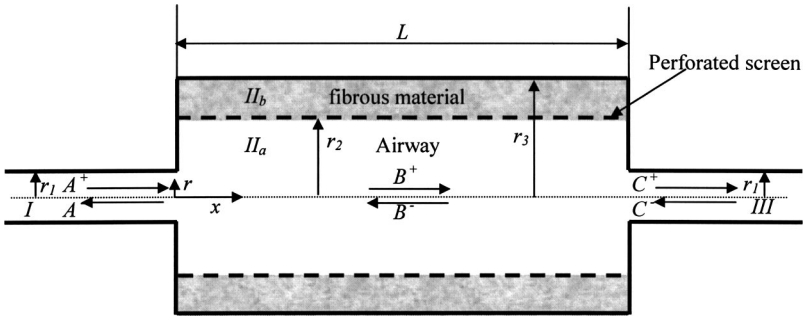


FIG. 1. Geometry of a dissipative silencer.

lindrical silencer. Following this introduction, Sec. II obtains the eigenvalues and eigenfunctions of the acoustic wave in the perforated absorbing chamber, and develops a two-dimensional analytical approach for the transmission loss by applying the pressure/velocity matching technique. Section III compares the analytical predictions with numerical results and experimental measurements, and discusses the effect of geometry, fiber resistance, and porosity on the transmission loss. The study is concluded with final remarks in Sec. IV.

II. ANALYTICAL APPROACH

Consider a cylindrical chamber of length L and radius r_3, with sound-absorbing material placed between radii r_2 and r_3, as shown in Fig. 1. The inlet and outlet pipes of radius r_1 are designated as domains I and III, and the chamber as II_a and II_b. A perforated screen with porosity ϕ is located at $r=r_2$ to separate the central airway (domain II_a) from the absorbing material (domain II_b) in the expansion chamber. The absorbing material is assumed to be homogeneous and isotropic, characterized by the complex speed of sound $\tilde{c}$ and density $\tilde{\rho}$ (The expressions are deferred to Appendix A).

A. Wave propagation in the inlet pipe (domain I)

For a two-dimensional axisymmetric wave propagation in the inlet duct (domain I), the solution of the Helmholtz equation is expressed as

$$P_A(r,x)=\sum_{n=0}^{\infty}(A_n^+e^{-jk_{x,A,n}x}+A_n^-e^{jk_{x,A,n}x})\psi_{A,n}(r), \quad (1)$$

with $j=\sqrt{-1}$, P_A being the acoustic pressure in the inlet pipe; A_n^+ and A_n^- the modal amplitudes corresponding to components traveling in the positive and negative x directions, respectively; $k_{x,A,n}$ the wave number in x direction with subscript x, A, n denoting axial direction, domain I, and order of the waves, respectively;

$$\psi_{A,n}(r)=J_0(k_{r,A,n}r), \quad (2)$$

is the eigenfunction for this circular duct with subscript r designating the radial direction and J_0 being the zeroth-order Bessel function of the first kind, $k_{r,A,n}$ the radial wave number satisfying the rigid wall boundary condition of

$$J_0'(k_{r,A,n}r_1)=0. \quad (3)$$

The axial wave number of the n mode is

$$k_{x,A,n}=\begin{cases}\sqrt{k_0^2-k_{r,A,n}^2}, & k_0>k_{r,A,n},\\ -j\sqrt{k_{r,A,n}^2-k_0^2}, & k_0<k_{r,A,n},\end{cases} \quad (4)$$

where $k_0=2\pi f/c_0$ is the wave number in air, f the frequency, and c_0 the speed of sound. In view of the linearized momentum equation, the particle velocity in the axial direction may then be written as

$$u_{x,A}(r,x)=\frac{1}{\rho_0\omega}\sum_{n=0}^{\infty}k_{x,A,n}\times(A_n^+e^{-jk_{x,A,n}x}-A_n^-e^{jk_{x,A,n}x})\psi_{A,n}(r), \quad (5)$$

where ρ_0 is the density of air and ω the angular velocity.

B. Wave propagation in the expansion chamber (domains II_a and II_b)

For the expansion chamber, the two-dimensional axisymmetric wave propagation in cylindrical coordinates (r,x) is governed by

$$\frac{\partial^2P}{\partial r^2}+\frac{1}{r}\frac{\partial P}{\partial r}+\frac{\partial^2P}{\partial x^2}+\kappa^2P=0, \quad (6)$$

where

$$\kappa=\begin{cases}k_0, & 0\leqslant r\leqslant r_2,\\ \tilde{k}, & r_2\leqslant r\leqslant r_3,\end{cases} \quad (7)$$

with $\tilde{k}=2\pi f/\tilde{c}$ denoting the wave number in the fibrous material.

The central airway and fiber material have the same axial wave number $k_{x,B,n}$ and different radial wave numbers $k_{r,B,n}$ (for the air) and $\tilde{k}_{r,B,n}$ (for the fiber) related by

$$k_{r,B,n}^2+k_{x,B,n}^2=k_0^2, \quad (8)$$

and

$$\tilde{k}_{r,B,n}^2+k_{x,B,n}^2=\tilde{k}^2. \quad (9)$$

The acoustic pressure is expressed as

$$P_B(r,x)=\sum_{n=0}^{\infty}(B_n^+e^{-jk_{x,B,n}x}+B_n^-e^{jk_{x,B,n}x})\psi_{B,n,P}(r), \quad (10)$$

where

$$P_B(r,x)=\begin{cases}P_{II_a}(r,x), & 0\leqslant r\leqslant r_2,\\ P_{II_b}(r,x), & r_2\leqslant r\leqslant r_3,\end{cases} \quad (11a)$$

$$\psi_{B,n,P}(r)=\begin{cases}\psi_{II_a,n,P}(r), & 0\leqslant r\leqslant r_2,\\ \psi_{II_b,n,P}(r), & r_2\leqslant r\leqslant r_3,\end{cases} \tag{11b}$$

with subscripts B, II_a, II_b, r denoting the chamber, domain II_a (the central airway), domain II_b (the fibrous material), and radial direction, respectively; B_n^+ and B_n^- the modal amplitudes; $\psi_{B,n,P}$ the transverse modal eigenfunction for the pressure. Substituting Eqs. (8)–(10) into (6) yields

$$\frac{\partial^2\psi_{B,n,P}(r)}{\partial r^2}+\frac{1}{r}\frac{\partial\psi_{B,n,P}(r)}{\partial r}+\kappa_{r,B,n}^2\psi_{B,n,P}(r)=0, \tag{12}$$

with

$$\kappa_{r,B,n}=\begin{cases}k_{r,B,n}, & 0\leqslant r\leqslant r_2,\\ \tilde{k}_{r,B,n}, & r_2\leqslant r\leqslant r_3.\end{cases} \tag{13}$$

The solution for Eq. (12) is expressed as

$$\psi_{B,n,P}(r)=\begin{cases}\psi_{II_a,n,P}(r)=C_1J_0(k_{r,B,n}r)+C_2Y_0(k_{r,B,n}r),\\ \qquad 0\leqslant r\leqslant r_2,\\ \psi_{II_b,n,P}(r)=C_3J_0(\tilde{k}_{r,B,n}r)+C_4Y_0(\tilde{k}_{r,B,n}r),\\ \qquad r_2\leqslant r\leqslant r_3,\end{cases} \tag{14}$$

where Y_0 denotes the zeroth-order Neumann function; C_1, C_2, C_3, and C_4 the coefficients related by the following four boundary conditions at $r=0$, r_2, and r_3:

(1) At $r=0$, the pressure is finite; thus, Eqs. (10), (11), and (14) yield

$$C_2=0. \tag{15}$$

(2) From Eqs. (10), (11), and (14), the rigid wall boundary condition at $r=r_3$ gives

$$u_{r,B}=\frac{j}{\tilde{\rho}\omega}\frac{\partial P_B}{\partial r}, \tag{16a}$$

$$u_{r,B}=0, \quad \text{for } r=r_3, \tag{16b}$$

$$C_3J_1(\tilde{k}_{r,B,n}r_3)+C_4Y_1(\tilde{k}_{r,B,n}r_3)=0, \tag{16c}$$

with $u_{r,B}$ being the radial particle velocity; J_1 and Y_1 the first-order Bessel and Neumann functions, respectively.

(3) From Eqs. (10), (11), and (14), the continuity of radial particle velocity u_r at $r=r_2$ yields

$$u_{r,II_a}=u_{r,II_b}, \quad \text{for } r=r_2, \tag{17a}$$

$$\frac{j}{\rho\omega}\frac{\partial P_{II_a}}{\partial r}=\frac{j}{\tilde{\rho}\omega}\frac{\partial P_{II_b}}{\partial r}, \quad \text{for } r=r_2, \tag{17b}$$

$$\frac{k_{r,B,n}}{\rho_0}[C_1J_1(k_{r,B,n}r_2)+C_2Y_1(k_{r,B,n}r_2)]=\frac{\tilde{k}_{r,B,n}}{\tilde{\rho}}[C_3J_1(\tilde{k}_{r,B,n}r_2)+C_4Y_1(\tilde{k}_{r,B,n}r_2)]. \tag{17c}$$

(4) The difference of acoustic pressure across the perforated screen ($r=r_2$) yields

$$P_{II_a}-P_{II_b}=\rho_0c_0\tilde{\zeta}_pu_{r,II_a}, \quad \text{for } r=r_2, \tag{18a}$$

$$[C_1J_0(k_{r,B,n}r_2)+C_2Y_0(k_{r,B,n}r_2)]-[C_3J_0(\tilde{k}_{r,B,n}r_2)+C_4Y_0(\tilde{k}_{r,B,n}r_2)]=\frac{\tilde{\zeta}_pk_{r,B,n}}{jk_0}[C_1J_1(k_{r,B,n}r_2)+C_2Y_1(k_{r,B,n}r_2)], \tag{18b}$$

with $\tilde{\zeta}_p$ (given in Appendix B) being the nondimensionalized acoustic impedance of the perforated screen.

By assuming

$$C_1=1, \tag{19}$$

Eqs. (15)–(18) yield

$$C_3=\frac{k_{r,B,n}\tilde{\rho}}{\tilde{k}_{r,B,n}\rho_0}\frac{J_1(k_{r,B,n}r_2)Y_1(\tilde{k}_{r,B,n}r_3)}{J_1(\tilde{k}_{r,B,n}r_2)Y_1(\tilde{k}_{r,B,n}r_3)-J_1(\tilde{k}_{r,B,n}r_3)Y_1(\tilde{k}_{r,B,n}r_2)}, \tag{20}$$

$$C_4=-\frac{J_1(\tilde{k}_{r,B,n}r_3)}{Y_1(\tilde{k}_{r,B,n}r_3)}C_3=-\frac{k_{r,B,n}\tilde{\rho}}{\tilde{k}_{r,B,n}\rho_0}\frac{J_1(k_{r,B,n}r_2)J_1(\tilde{k}_{r,B,n}r_3)}{J_1(\tilde{k}_{r,B,n}r_2)Y_1(\tilde{k}_{r,B,n}r_3)-J_1(\tilde{k}_{r,B,n}r_3)Y_1(\tilde{k}_{r,B,n}r_2)}, \tag{21}$$

and the characteristic equation

$$\frac{\rho_0\tilde{k}_{r,B,n}}{\tilde{\rho}k_{r,B,n}}\left[\frac{J_0(k_{r,B,n}r_2)}{J_1(k_{r,B,n}r_2)}+\frac{j\tilde{\zeta}_pk_{r,B,n}}{k_0}\right]=\frac{J_0(\tilde{k}_{r,B,n}r_2)Y_1(\tilde{k}_{r,B,n}r_3)-Y_0(\tilde{k}_{r,B,n}r_2)J_1(\tilde{k}_{r,B,n}r_3)}{J_1(\tilde{k}_{r,B,n}r_2)Y_1(\tilde{k}_{r,B,n}r_3)-Y_1(\tilde{k}_{r,B,n}r_2)J_1(\tilde{k}_{r,B,n}r_3)}. \tag{22}$$

In view of Eqs. (8) and (9), Eq. (22) yields the solutions for the axial wave number $k_{x,B,n}$ for a given frequency. (The details of the solution technique are deferred to Appendix C.) The transverse eigenfunctions for the pressure can then be determined as

$$\psi_{B,n,P}(r)=\begin{cases}J_0(k_{r,B,n}r), & 0\leqslant r\leqslant r_2,\\ C_5\left[J_0(\tilde{k}_{r,B,n}r)-\dfrac{J_1(\tilde{k}_{r,B,n}r_3)}{Y_1(\tilde{k}_{r,B,n}r_3)}Y_0(\tilde{k}_{r,B,n}r)\right], & r_2\leqslant r\leqslant r_3,\end{cases} \tag{23}$$

with

$$C_5=\left[J_0(k_{r,B,n}r_2)+\frac{j\tilde{\zeta}_p k_{r,B,n}}{k_0}J_1(k_{r,B,n}r_2)\right]\frac{Y_1(\tilde{k}_{r,B,n}r_3)}{J_0(\tilde{k}_{r,B,n}r_2)Y_1(\tilde{k}_{r,B,n}r_3)-J_1(\tilde{k}_{r,B,n}r_3)Y_0(\tilde{k}_{r,B,n}r_2)}.$$

Equations (10), (11), and (23) yield

$$P_B(r)=\begin{cases}\sum_{n=0}^{\infty}(B_n^+e^{-jk_{x,B,n}x}+B_n^-e^{jk_{x,B,n}x})J_0(k_{r,B,n}r), & 0\leqslant r\leqslant r_2,\\ \sum_{n=0}^{\infty}(B_n^+e^{-jk_{x,B,n}x}+B_n^-e^{jk_{x,B,n}x})C_5\left[J_0(\tilde{k}_{r,B,n}r)-\frac{J_1(\tilde{k}_{r,B,n}r_3)}{Y_1(\tilde{k}_{r,B,n}r_3)}Y_0(\tilde{k}_{r,B,n}r)\right], & r_2\leqslant r\leqslant r_3.\end{cases} \tag{24}$$

From the linearized momentum equation

$$u_{x,B}=\begin{cases}\frac{j}{\rho_0\omega}\frac{\partial P_{II_a}}{\partial x}, & 0\leqslant r\leqslant r_2,\\ \frac{j}{\tilde{\rho}\omega}\frac{\partial P_{II_b}}{\partial x}, & r_2\leqslant r\leqslant r_3,\end{cases}$$

the particle velocity in the axial direction is then obtained as

$$u_{x,B}(r)=\frac{1}{\rho_0\omega}\sum_{n=0}^{\infty}k_{x,B,n}\times(B_n^+e^{-jk_{x,B,n}x}-B_n^-e^{jk_{x,B,n}x})\psi_{B,n,u_x}(r), \tag{25a}$$

with

$$\psi_{B,n,u_x}(r)=\begin{cases}J_0(k_{r,B,n}r), & 0\leqslant r\leqslant r_2,\\ C_5\frac{\rho_0}{\tilde{\rho}}\left[J_0(\tilde{k}_{r,B,n}r)-\frac{J_1(\tilde{k}_{r,B,n}r_3)}{Y_1(\tilde{k}_{r,B,n}r_3)}Y_0(\tilde{k}_{r,B,n}r)\right], & r_2\leqslant r\leqslant r_3,\end{cases} \tag{25b}$$

being the transverse modal eigenfunction for the velocity.

C. Wave propagation in the outlet pipe (domain III)

The acoustic pressure and axial velocity in the outlet pipe (domain III) are similar to those in the inlet pipe (domain I) and expressed as

$$P_C(r,x)=\sum_{n=0}^{\infty}(C_n^+e^{-jk_{x,C,n}(x-L)}+C_n^-e^{jk_{x,C,n}(x-L)})\psi_{C,n}(r), \tag{26a}$$

and

$$u_{x,C}(r,x)=\frac{1}{\rho_0\omega}\sum_{n=0}^{\infty}k_{x,C,n}(C_n^+e^{-jk_{x,C,n}(x-L)}-C_n^-e^{jk_{x,C,n}(x-L)})\psi_{C,n}(r), \tag{26b}$$

where the subscript C denotes domain III, C_n^+ and C_n^- are the amplitudes; with both the eigenfunction $\psi_{C,n}(r)$ and wave number $k_{x,C,n}$ being similar to those in the inlet pipe.

D. Transmission loss prediction

With the expressions of pressure and particle velocity of the inlet, outlet, and expansion chamber [Eqs. (1), (5), and (24–26)], transmission loss can then be obtained by solving the unknown coefficients A_n, B_n, and C_n using the boundary conditions at the expansion ($x=0$) and contraction ($x=L$). At the interfaces of the expansion and contraction, the acoustic pressure and velocity continuity conditions reveal

$$P_A=P_B, \quad \text{for } 0\leqslant r\leqslant r_1, \ x=0, \tag{27a}$$

$$u_{x,B}=\begin{cases}u_{x,A}, & \text{for } 0\leqslant r\leqslant r_1, \ x=0,\\ 0, & \text{for } r_1\leqslant r\leqslant r_3, \ x=0,\end{cases} \tag{27b}$$

$$P_C=P_B, \quad \text{for } 0\leqslant r\leqslant r_1, \ x=L, \tag{27c}$$

$$u_{x,B}=\begin{cases}u_{x,C}, & \text{for } 0\leqslant r\leqslant r_1, \ x=L,\\ 0, & \text{for } r_1\leqslant r\leqslant r_3, \ x=L.\end{cases} \tag{27d}$$

In view of the expressions of the pressure and velocity as infinite series of unknown amplitudes in Eqs. (1), (5), and (24–26), Eq. (27) gives

$$\sum_{n=0}^{\infty}(A_n^++A_n^-)\psi_{A,n}(r)=\sum_{n=0}^{\infty}(B_n^++B_n^-)\psi_{B,n,P}(r), \quad \text{for } 0\leqslant r\leqslant r_1, \tag{28a}$$

$$\sum_{n=0}^{\infty}k_{x,B,n}(B_n^+-B_n^-)\psi_{B,n,u_x}(r)=\begin{cases}\sum_{n=0}^{\infty}k_{x,A,n}(A_n^+-A_n^-)\psi_{A,n}(r), & \text{for } 0\leqslant r\leqslant r_1,\\ 0, & \text{for } r_1\leqslant r\leqslant r_3,\end{cases} \tag{28b}$$

$$\sum_{n=0}^{\infty}(C_n^++C_n^-)\psi_{C,n}(r)=\sum_{n=0}^{\infty}(B_n^+e^{-jk_{x,B,n}L}+B_n^-e^{jk_{x,B,n}L})\psi_{B,n,P}(r), \quad \text{for } 0\leqslant r\leqslant r_1, \tag{28c}$$

$$\sum_{n=0}^{\infty} k_{x,B,n}(B_n^+ e^{-jk_{x,B,n}L} - B_n^- e^{jk_{x,B,n}L})\psi_{B,n,u_x}(r)$$

$$= \begin{cases} \sum_{n=0}^{\infty} k_{x,C,n}(C_n^+ - C_n^-)\psi_{C,n}(r), & \text{for } 0 \leqslant r \leqslant r_1, \\ 0, & \text{for } r_1 \leqslant r \leqslant r_3. \end{cases} \tag{28d}$$

In Eq. (28), the infinite series of unknown amplitudes need to be truncated to a suitable number. Then, the same number of equations is solved for the amplitudes of acoustic waves. An approach proposed in Ref. 18 is adopted here to match the sound field. Imposing the continuities of the integral of the pressure and axial velocity over discrete zones of the interfaces at the expansion ($x=0$) and contraction ($x=L$), Eq. (28) yields the pressure and velocity matching conditions as

$$\sum_{n=0}^{N} (A_n^+ + A_n^-) \int_0^{r_{m,P}} \psi_{A,n}(r)dr = \sum_{n=0}^{N} (B_n^+ + B_n^-) \int_0^{r_{m,P}} \psi_{B,n,P}(r)dr, \tag{29a}$$

$$\sum_{n=0}^{N} k_{x,B,n}(B_n^+ - B_n^-) \int_0^{r_{m,u}} \psi_{B,n,u_x}(r)dr = \begin{cases} \sum_{n=0}^{N} k_{x,A,n}(A_n^+ - A_n^-) \int_0^{r_{m,u}} \psi_{A,n}(r)dr, \\ \quad \text{for } 0 \leqslant r_{m,u} \leqslant r_1, \\ \sum_{n=0}^{N} k_{x,A,n}(A_n^+ - A_n^-) \int_0^{r_1} \psi_{A,n}(r)dr, \\ \quad \text{for } r_1 \leqslant r_{m,u} \leqslant r_3, \end{cases} \tag{29b}$$

$$\sum_{n=0}^{N} (C_n^+ + C_n^-) \int_0^{r_{m,P}} \psi_{C,n}(r)dr = \sum_{n=0}^{N} (B_n^+ e^{-jk_{x,B,n}L} + B_n^- e^{jk_{x,B,n}L}) \int_0^{r_{m,P}} \psi_{B,n,P}(r)dr, \tag{29c}$$

$$\sum_{n=0}^{N} k_{x,B,n}(B_n^+ e^{-jk_{x,B,n}L} - B_n^- e^{jk_{x,B,n}L}) \int_0^{r_{m,u}} \psi_{B,n,u_x}(r)dr = \begin{cases} \sum_{n=0}^{N} k_{x,C,n}(C_n^+ - C_n^-) \int_0^{r_{m,u}} \psi_{C,n}(r)dr, \\ \quad \text{for } 0 \leqslant r_{m,u} \leqslant r_1, \\ \sum_{n=0}^{N} k_{x,C,n}(C_n^+ - C_n^-) \int_0^{r_1} \psi_{C,n}(r)dr, \\ \quad \text{for } r_1 \leqslant r_{m,u} \leqslant r_3, \end{cases} \tag{29d}$$

with

$$r_{m,P} = \frac{m}{N+1} r_1, \quad m = 1,...,N+1; \tag{30a}$$

$$r_{m,u} = \frac{m}{N+1} r_3, \quad m = 1,...,N+1. \tag{30b}$$

In view of $4(N+1)$ coefficients solved in Eq. (29) and the assumptions that: (1) the incoming wave being planar and A_0^+ be unity; (2) an anechoic termination be imposed at the exit by setting C_n^- to zero; and (3) all transmitted waves in the outlet pipe are nonpropagating modes except the first mode with C_0^+, the transmission loss is then determined as

$$\mathrm{TL} = -20 \log_{10} |C_0^+|. \tag{31}$$

III. RESULTS AND DISCUSSION

A single-pass concentric silencer with acoustic absorbing material and uniformly perforated duct is studied here (recall Fig. 1). Parameters of the basic model are given as $L=25.72$ cm, $r_1=2.45$ cm, $r_2=2.45$ cm, $r_3=8.22$ cm, $R=4896$ Rayls/m (for a filling density of 100 g/l), duct porosity $\phi=8\%$, thickness of the perforated duct $t_w=0.09$ cm, and diameter of the holes $d_h=0.249$ cm. The two-microphone technique and BEM[4] are also used to obtain the transmission loss experimentally and computationally.

To check the accuracy of analytical approach for the basic model, Fig. 2 presents the transmission loss determined from the analytical approach (with different numbers of higher order modes), 3D BEM, and the experiments. For the current geometry and frequency range of interest, the analytical results are accurate enough for $N \geqslant 9$, which are essentially identical to those for $N=12$ to the degree that the latter cannot be distinguished in the figure. These results agree well with the BEM predictions and show a reasonable comparison with the measurements.

Analytical results for absorbing silencers with different flow resistivities ($R=1000$, 4896, and 17378 Rayls/m) are presented in Fig. 3, in contrast to an empty chamber. Higher flow resistivity tends to improve the performance of the silencer at medium to high frequencies, while moving the transmission loss peak to lower frequencies. At low frequencies, the transmission loss is reduced with flow resistivity. The acoustic performance of the silencer with $R=4896$ Rayls/m is better than the one with no absorbing material nearly at all frequencies, however, the transmission loss of the silencer with $R=17\,378$ Rayls/m deteriorates slightly below 250 Hz.

Analytical results for silencers with different perforation porosities ($\phi=2\%$, 8%, and 50%) are presented in Fig. 4, including the limiting case of removed screen. At relatively high frequencies, the attenuation is significantly improved with porosity, while shifting the peak to higher frequencies. Higher perforations, however, tend to yield somewhat lower TL at low frequencies. The silencer with $\phi=2\%$ exhibits a behavior similar to that of a Helmholtz resonator, with its resonance frequency around 700 Hz and a rapidly decreasing transmission loss with frequency. The performance of the silencer with $\phi=50\%$ is close to the one without the perforated screen, as expected.

The analytical results with varying fiber thickness ($r_2=2.45$, 3.45, and 4.45 cm) are depicted in Fig. 5. Similar to a dissipative expansion chamber without perforated screen,[18]

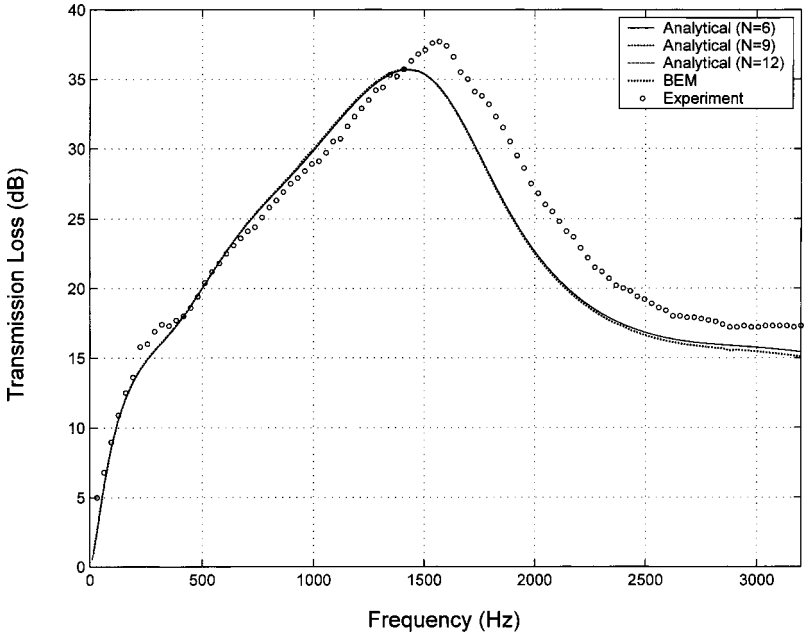


FIG. 2. Transmission loss of a single-pass perforated absorbing silencer ($L=25.72$ cm, $r_1=2.45$ cm, $r_2=2.45$ cm, $r_3=8.22$ cm, $R=4896$ Rayls/m, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm).

increasing the fiber thickness improves the attenuation particularly at mid to high frequencies. The location of peak attenuation is, however, almost maintained.

Figure 6 shows the effect of outer chamber radius ($r_3=5.055$, 8.22, and 10 cm). Increasing r_3 improves the performance of the absorbing silencers at low frequencies, while shifting the peak attenuation to lower frequencies. Unlike reactive chambers, however, the larger r_3 does not necessarily lead to higher attenuation peaks. Figure 7 illustrates the effect of chamber length ($L=15$, 25.72, and 40 cm) on the transmission loss of absorbing silencers. As expected, increasing L improves the attenuation at all frequencies.

IV. CONCLUSIONS

A two-dimensional, closed-form analytical approach has been developed to predict the acoustic attenuation of a perforated single-pass, concentric cylindrical expansion chamber filled with fibrous material. By using the boundary conditions at the central, perforated screen, and rigid wall, the governing eigenequation is obtained, which determines the sound field in the dissipative chamber. With the eigenvalues and eigenfunctions solved from the eigenequation, the transmission loss is predicted by applying the pressure and particle velocity matching at the interfaces of the expansion and contraction. The analytical predictions show a reasonable agreement with the experimental and BEM results for the frequency range of interest. The present study also illustrates the effects of chamber geometry, fiber parameters, and porosity on the acoustic attenuation of perforated dissipative silencers.

APPENDIX A: ACOUSTICAL PROPERTIES OF FIBER MATERIAL

The absorption of acoustic waves in fibrous material is mainly due to viscous dissipation, which involves complex-

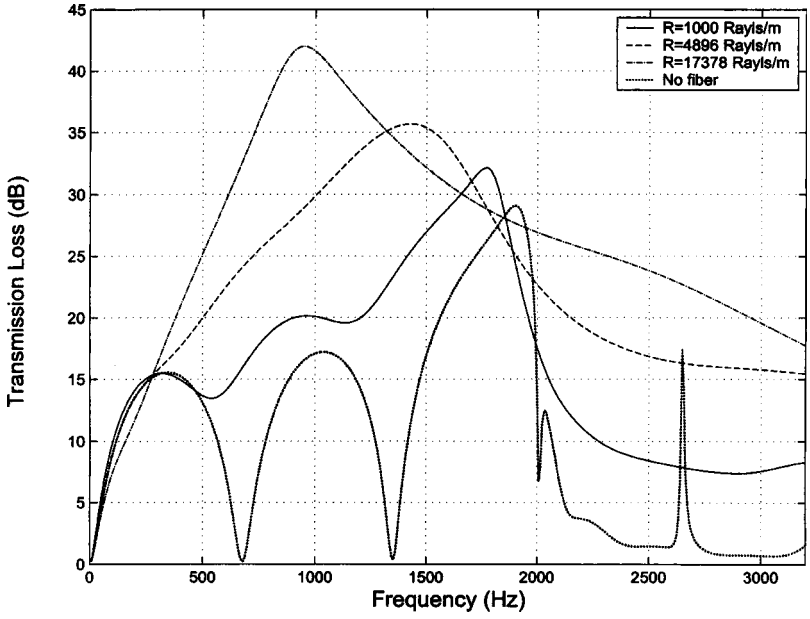


FIG. 3. Transmission loss of a single-pass perforated absorbing silencer with different flow resistivities ($L=25.72$ cm, $r_1=2.45$ cm, $r_2=2.45$ cm, $r_3=8.22$ cm, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm).

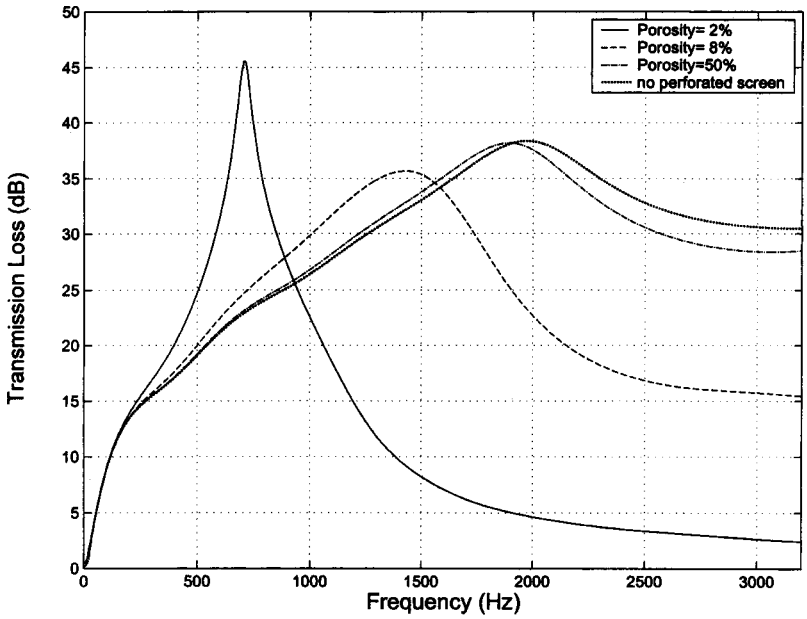


FIG. 4. Transmission loss of a single-pass perforated absorbing silencer with different porosites ($L=25.72$ cm, $r_1=2.45$ cm, $r_2=2.45$ cm, $r_3=8.22$ cm, $R=4896$ Rayls/m, $d_h=0.249$ cm, and $t_w=0.09$ cm).

valued characteristic impedance $\tilde{Z}=\tilde{\rho}\tilde{c}$ and wave number $\tilde{k}=2\pi f/\tilde{c}$. For the absorbing material in the present study, $\tilde{Z}$ and $\tilde{k}$ are given as[21]

$$\frac{\tilde{Z}}{Z_0}=[1+0.0855(f/R)^{-0.754}]+j[-0.0765(f/R)^{-0.732}], \tag{A1}$$

$$\frac{\tilde{k}}{k}=[1+0.1472(f/R)^{-0.577}]+j[-0.1734(f/R)^{-0.595}], \tag{A2}$$

where $Z_0=\rho_0 c_0$ is the characteristic impedance of the air; R [mks Rayls/m] denotes the flow resistivity. The measured flow resistivities R are 4896 and 17 378 Rayls/m, corresponding to material densities of 100 and 200 g/l, respectively.

APPENDIX B: ACOUSTIC IMPEDANCE OF PERFORATES

In Eq. (18), the nondimensionalized perforate acoustic impedance $\tilde{\zeta}_p$ relates the acoustic pressure in the inner duct and outer chamber through the interface. Sullivan and Crocker[22] presented an empirical expression for the acoustic impedance of perforate holes as

$$\tilde{\zeta}_p=[0.006+jk_0(t_w+0.75d_h)]/\phi. \tag{B1}$$

In the presence of fiber (for perforations facing absorbing material), Eq. (B1) has been modified by Selamet *et al.*[4] as

$$\tilde{\zeta}_p=\left[0.006+jk_0\left\{t_w+0.375d_h\left(1+\frac{\tilde{Z}}{Z_0}\frac{\tilde{k}}{k_0}\right)\right\}\right]\Big/\ \phi, \tag{B2}$$

which is also used in this study.

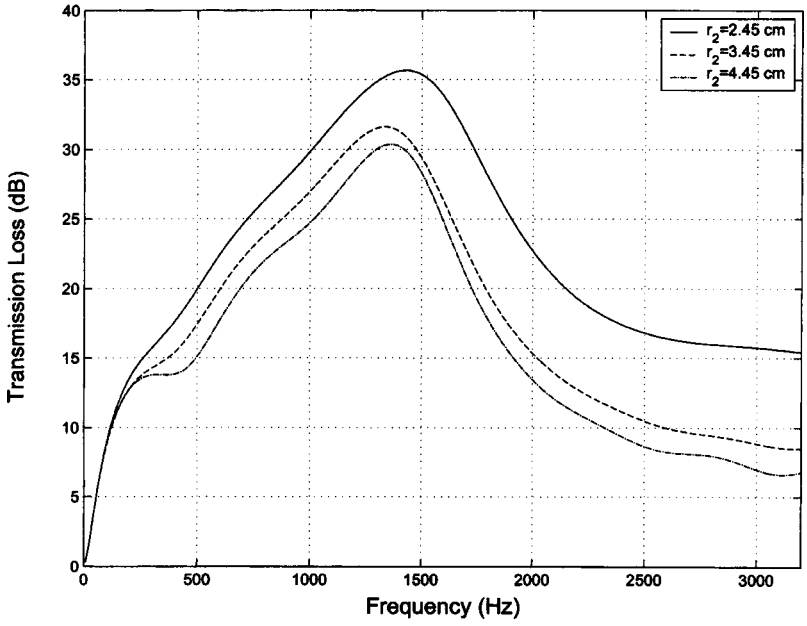


FIG. 5. Transmission loss of a single-pass perforated absorbing silencer with different central airway radii ($L=25.72$ cm, $r_1=2.45$ cm, $r_3=8.22$ cm, $R=4896$ Rayls/m, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm).

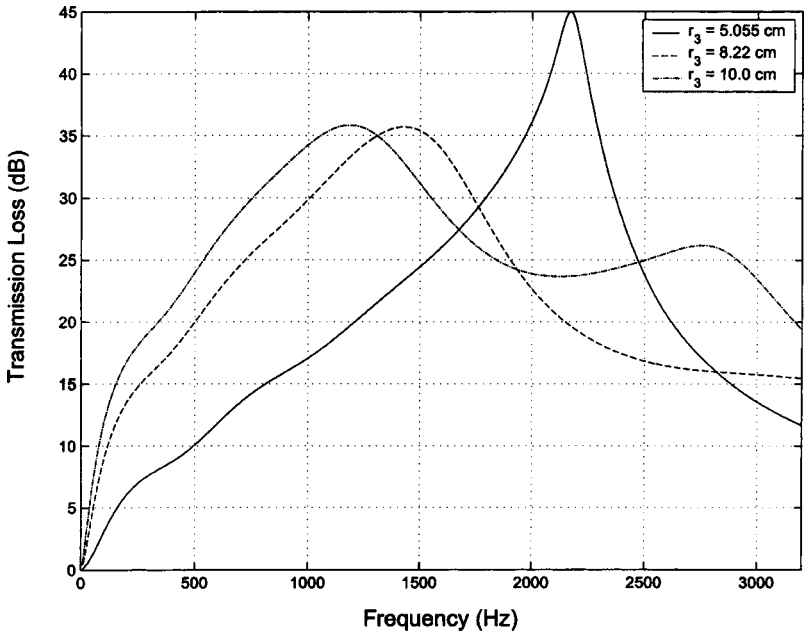


FIG. 6. Transmission loss of a single-pass perforated absorbing silencer with different chamber radii ($L=25.72$ cm, $r_1=2.45$ cm, $r_2=2.45$ cm, $R=4896$ Rayls/m, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm).

APPENDIX C: SOLUTION OF THE CHARACTERISTIC EQUATION

The axial wave numbers are obtained from the solution of

$$F(R,\tilde{\zeta}_p,k_x,f)$$
$$=\rho_0\tilde{k}_r[J_1(\tilde{k}_r r_2)Y_1(\tilde{k}_r r_3)-Y_1(\tilde{k}_r r_2)J_1(\tilde{k}_r r_3)]$$
$$\times[k_0 J_0(k_r r_2)+j\tilde{\zeta}_p k_r J_1(k_r r_2)]-\tilde{\rho}k_r k_0 J_1(k_r r_2)$$
$$\times[J_0(\tilde{k}_r r_2)Y_1(\tilde{k}_r r_3)-Y_0(\tilde{k}_r r_2)J_1(\tilde{k}_r r_3)]=0, \quad \text{(C1)}$$

where

$$k_r^2+k_x^2=k_0^2, \quad \tilde{k}_r^2+k_x^2=\tilde{k}^2.$$

Given an initial approximation for $k_{x,0}$ in the secant method used here, the root of Eq. (C1) is obtained by the loop

$$k_{x,i+1}=k_{x,i}-F(R,\tilde{\zeta}_p,k_{x,i},f)$$
$$\times\frac{(1+\Delta)k_{x,i}-(1-\Delta)k_{x,i}}{F(R,\tilde{\zeta}_p,(1+\Delta)k_{x,i},f)-F(R,\tilde{\zeta}_p,(1-\Delta)k_{x,i},f)}, \quad \text{(C2)}$$

where $i=0,1,2,...,i_{\max}$, Δ is a small constant, and $i_{\max}$ is an integer designating the loop number. The selection of a suitable initial guesstimate for the desired axial wave number $k_{x,0}$ is critical in this approach. For example, Cummings[10,11] obtained the roots of a dissipative silencer by the secant method with initial values chosen from a rectangular grid in the complex plane, which may cause "jumping" of eigenvalues. The present work avoids such potential eigenvalue jumping by determining the initial values based on a method

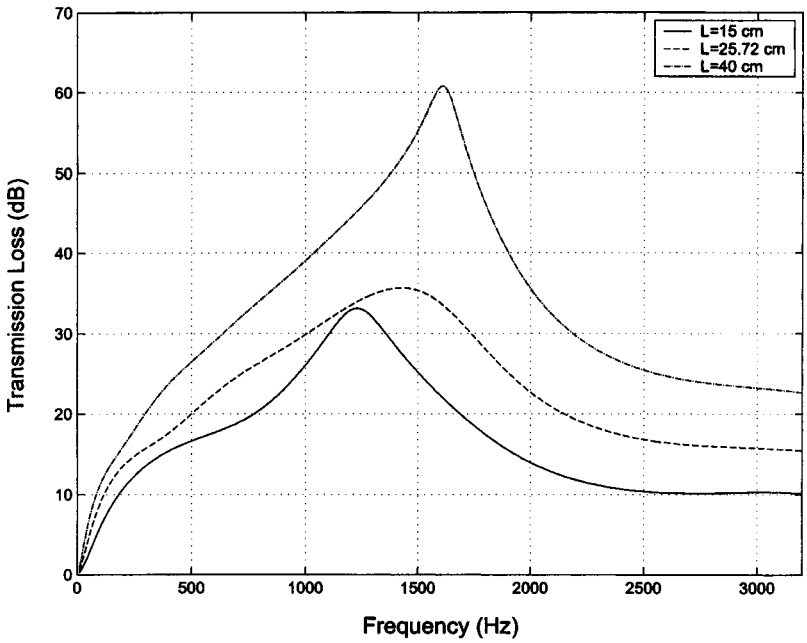


FIG. 7. Transmission loss of a single-pass perforated absorbing silencer with different chamber lengths ($r_1=2.45$ cm, $r_2=2.45$ cm, $r_3=8.22$ cm, $R=4896$ Rayls/m, $\phi=8\%$, $d_h=0.249$ cm, and $t_w=0.09$ cm).

originated from the solution of dispersion equations for shell–fluid coupled systems.[23–27]

Since the characteristic impedance and wave number in fibrous materials approach the corresponding values in air at very high frequencies, the roots of Eq. (C1) are obtained first at the upper-frequency limit of the range of interest, $f_{\max}$. Let $R=0$ and $\tilde{\zeta}_p=0$ in $F(R,\tilde{\zeta}_p,k_x,f_{\max})$ of Eq. (C1). The solution of $F(0,0,k_x,f_{\max})=0$ then yields the axial wave numbers of a duct without fiber or perforation at $f_{\max}$: $k_x^1(0,0,f_{\max})$, $k_x^2(0,0,f_{\max})$,..., with the superscripts denoting the mode order. With the foregoing initial approximation, the roots of the characteristic equation

$$F\left(\frac{R}{N_{R,\tilde{\zeta}_p}},\frac{\tilde{\zeta}_p}{N_{R,\tilde{\zeta}_p}},k_x,f_{\max}\right)=0 \tag{C3}$$

(with $N_{R,\tilde{\zeta}_p}$ being a large integer), can be readily obtained from the secant method as $k_x^1[(R/N_{R,\tilde{\zeta}_p}),(\tilde{\zeta}_p/N_{R,\tilde{\zeta}_p}),f_{\max}]$, $k_x^2[(R/N_{R,\tilde{\zeta}_p}),(\tilde{\zeta}_p/N_{R,\tilde{\zeta}_p}),f_{\max}]$,.... Since the axial wave numbers of Eq. (C3) are very close to those of $F(0,0,k_x,f_{\max})=0$, the secant method with such initial values is proved to be successful and avoids eigenvalue jumping. For $n_{R,\tilde{\zeta}_p}=1,2,...,N_{R,\tilde{\zeta}_p}-1$, the roots of $F([(n_{R,\tilde{\zeta}_p}+1)/N_{R,\tilde{\zeta}_p}]R,[(n_{R,\tilde{\zeta}_p}+1)/N_{R,\tilde{\zeta}_p}]\tilde{\zeta}_p,k_x,f_{\max})=0$ are determined with the initial values estimated as the solutions of $F[(n_{R,\tilde{\zeta}_p}/N_{R,\tilde{\zeta}_p})R,(n_{R,\tilde{\zeta}_p}/N_{R,\tilde{\zeta}_p})\tilde{\zeta}_p,k_x,f_{\max}]=0$. Thus, for the dissipative expansion chamber with flow resistivity R and perforation $\tilde{\zeta}_p$, the roots of the characteristic equation $F(R,\tilde{\zeta}_p,x,f_{\max})=0$ have been obtained at the highest frequency of interest $f_{\max}$. With the results now available at $f_{\max}$, the axial wave numbers can then be determined at all frequencies by a similar secant method. The frequency is reduced by a suitable step Δf ($=10$ Hz, for example), and the solution for the previous frequency is used as the initial values at the new frequency $f_{\max}-\Delta f$. The characteristic equation is thus solved at all frequencies, leading to the dispersion curves for the dissipative expansion silencer. This method is commonly adopted to get the dispersion curves for cylindrical shell–fluid coupled systems (both fluid-filled shell[23–27] and the shell immersed in fluid[26]). Numerical study demonstrates that this method avoids the eigenvalue jumping addressed in Refs. 10 and 11. In addition, the present approach can save significant computational time as compared to the numerical methods (BEM/FEM). For the basic model with $r_1=2.45$ cm, $r_2=2.45$ cm, $r_3=8.22$ cm, $R=4896$ Rayls/m, and $\phi=8\%$, the computational time for the present analytical method is much less than that of BEM.

[1] K. S. Peat and K. L. Rathi, "A finite element analysis of the convected acoustic wave motion in dissipative silencers," J. Sound Vib. **184**, 529–545 (1995).

[2] R. J. Astley and A. Cummings, "A finite element scheme for attenuation in ducts lined with porous material: Comparison with experiment," J. Sound Vib. **116**, 239–263 (1987).

[3] A. F. Seybert, R. A. Seman, and M. D. Lattuca, "Boundary element prediction of sound propagation in ducts containing bulk absorbing materials," J. Vib. Acoust. **120**, 976–981 (1998).

[4] A. Selamet, I. J. Lee, Z. L. Ji, and N. T. Huff, "Acoustic attenuation performance of perforated absorbing silencers, SAE Noise and Vibration Conference and Exposition," SAE Paper No. 2001-01-1435, Traverse City, MI, 30 April–3 May (2001).

[5] A. Selamet, I. J. Lee, and N. T. Huff, "Acoustic attenuation of hybrid silencers," J. Sound Vib. **262**, 509–527 (2003).

[6] SYSNOISE users' manual, Rev. 5.5, LMS International, 2001.

[7] R. A. Scott, "The propagation of sound between walls of porous material," Proc. Phys. Soc. **58**, 358–368 (1946).

[8] S. H. Ko, "Theoretical analyses of sound attenuation in acoustically lined flow ducts separated by porous splitters (rectangular, annular, and circular ducts)," J. Sound Vib. **39**, 471–487 (1975).

[9] B. Nilsson and O. Brander, "The propagation of sound in cylindrical ducts with mean flow and bulk reacting lining. I. Modes in an infinite duct," IMA J. Appl. Math. **26**, 269–298 (1980).

[10] A. Cummings and I. J. Chang, "Internal mean flow effects on the characteristics of bulk-reacting liners in circular ducts," Acustica **64**, 169–178 (1987).

[11] A. Cummings and I. J. Chang, "Sound attenuation of a finite length dissipative flow duct silencer with internal mean flow in the absorbent," J. Sound Vib. **127**, 1–17 (1988).

[12] K. S. Peat, "A transfer matrix for an absorption silencer element," J. Sound Vib. **146**, 353–360 (1991).

[13] R. Kirby, "Simplified techniques for predicting the transmission loss of a circular dissipative silencer," J. Sound Vib. **243**, 403–426 (2001).

[14] C. N. Wang, "Numerical decoupling analysis of a resonator with absorbent material," Appl. Acoust. **58**, 109–122 (1999).

[15] M. L. Munjal and P. T. Thawani, "Effect of protective layer on the performance of absorptive ducts," Noise Control Eng. J. **45**, 14–18 (1997).

[16] M. L. Munjal, "Analysis and design of pod silencers," J. Sound Vib. **262**, 497–507 (2003).

[17] M. B. Xu, X. M. Zhang, and W. H. Zhang, "The effect of wall joint on the vibrational power flow propagation in a fluid-filled shell," J. Sound Vib. **224**, 395–410 (1999).

[18] M. B. Xu, A. Selamet, I. J. Lee, and N. T. Huff, "Sound attenuation in dissipative expansion chambers," J. Sound Vib. (in press).

[19] U. Ingard and R. H. Bolt, "Absorption characteristics of acoustic material with perforated facings," J. Acoust. Soc. Am. **23**, 533–540 (1951).

[20] R. Kirby and A. Cummings, "The impedance of perforated plates subjected to grazing gas flow and backed by porous media," J. Sound Vib. **217**, 619–636 (1998).

[21] M. Nice, Owens Corning Automotive, Internal report (1999).

[22] J. W. Sullivan and M. J. Crocker, "Analysis of concentric-tube resonators having unpartitioned cavities," J. Acoust. Soc. Am. **64**, 207–215 (1978).

[23] C. R. Fuller and F. J. Fahy, "Characteristics of wave propagation and energy distribution in cylindrical elastic shells filled with fluid," J. Sound Vib. **81**, 501–518 (1981).

[24] M. B. Xu, "Wave propagation and control in pipes conveying fluid," MS thesis, Huazhong University of Science and Technology (1995).

[25] M. B. Xu, "Vibro-acoustic properties of elastic pipes filled with convected fluid," Appl. Acoust. (submitted).

[26] M. B. Xu, "Wave propagation and power flow in cylindrical shell—fluid coupled systems," Ph.D. dissertation, Huazhong University of Science and Technology (1999).

[27] M. B. Xu, "Three methods for analyzing forced vibration of a fluid-filled cylindrical shell," Appl. Acoust. **64**, 731–752 (2003).

Tunneling effect in sound transmission loss determination: Theoretical approach

Bong-Ki Kim,[a] Hyun-Ju Kang, Jae-Seung Kim, Hyun-Sil Kim, and Sang-Ryul Kim
Acoustics Group, Korea Institute of Machinery and Materials (KIMM), 101 P.O. Box, Yusung, Daejon 305-600, Korea

(Received 14 June 2002; revised 7 February 2004; accepted 12 February 2004)

The aim of this study is to numerically evaluate a tunneling effect in the laboratory measurement of sound transmission loss. The tunneling effect arises from the depth of an aperture in the common wall between the source and receiving rooms. Variations of the sound transmission loss with the parameters of panel location, tunnel depth, and panel size are investigated. The difference in sound transmission loss is quite evident below the coincidence frequency and it greatly depends on the panel location in the tunnel. In comparison with the transmission loss of a finite plate in an infinite rigid baffle (with no tunnel) the maximum difference occurs in the laboratory measurement when the panel is placed at the center of the tunnel, while a better estimation of true transmission loss is obtained when the panel is located at either end. The results provide an added guideline for the standard laboratory test method for sound transmission loss. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1698815]

PACS numbers: 43.55.Rg, 43.55.Nd, 43.55.Ti [SFW] Pages: 2100–2109

I. INTRODUCTION

The sound insulation performance of a panel can be generally characterized by means of measurements of sound transmission loss (STL) in two reverberation chambers having a common aperture. A considerable effort has been devoted to match theoretical predictions with experimental measurements in reverberation chambers to establish the valid theoretical model.[1–4] From the standpoint of experimental method, reproducibility[5] in relation to consistency and accuracy of sound insulation measurements has been an important issue.

The majority of studies on reproducibility have been devoted to experimental investigations involving round robin tests.[6–8] The results indicate that different STLs from different laboratories were obtained even though the same panels were used for comparison.[7,9–11] Such large discrepancies in reproducibility have seriously deteriorated the reliability of STL measurements, thus making it difficult to establish a valid analytical model for sound insulation.

Among the various reasons for the large deviations in reproducibility tests, an evident factor experimentally proven is the so-called "tunneling effect."[12] It is well known that when measuring STL in a laboratory, the location of the specimen in an aperture obviously affects the results due to the tunneling effect. This is due to the fact that a tunnel with a depth comparable to the acoustic wavelength behaves like an acoustic duct, thus altering the sound field of both sides of the panel, presumably resulting in a nondiffuse sound field. Several experimental studies[12–14] reported that the values of STL are critically dependent on the location of the panel in the tunnel as shown in Fig. 1. For example, Halliwell *et al.*[12] showed that the case of a center-located panel yielded a lower STL than that of one flush with the end of tunnel.

Although a number of experiments have been performed, a theoretical study of the tunneling effect has not yet been reported. Furthermore, there is a limit to fully explain the tunneling effect by experiments because the results also include other test conditions such as repeatability, panel size, diffuseness, mounting conditions of the test panel, room volume, and room shape.[6] To better understand the tunneling effect, it will be necessary to develop a theoretical model for calculating the STL of a finite panel with the tunnel.

Basically, a tunnel can modify the STL in several ways. The pressure field of the tunnel alters the structural response of the panel, giving rise to a feedback loop coupling the elastic and acoustic problems. The analysis of a fluid loaded finite plate is significantly more complex than the corresponding problem involving an infinite plate.[15,16] The general procedure for solving this kind of problem is to use the Helmholtz integral formula[16] for expressing the radiated pressure and matching boundary conditions between plate vibration and the acoustic field of the tunnel which are expanded in terms of the corresponding normal modes. The Fourier transform method[17,18] can be used as an alternative to the Helmholtz integral formula.

This paper deals with a two-dimensional system and is concerned with the STL of a finite panel, installed in a tunnel between the source and receiving room. Thus, the panel can be considered as an infinite strip with a finite height. In the following, the scattering and radiation fields resulting from the tunnel are expressed by the Fourier transform. To obtain the vibration response of the plate as well as the transmitted acoustic power, it is necessary to use an inverse transform in the wavenumber domain. Since the required inversion cannot be evaluated in a closed form, the technique of contour integration is adopted. The method is applied to a finite glass panel and the effects of tunnel depth, panel location, and panel height are illustrated and the STL values are compared for the cases with and without tunnels.

[a]Electronic mail: bkkim@kimm.re.kr

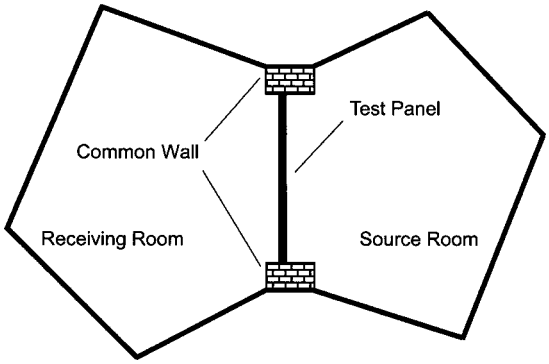


FIG. 1. Schematic illustration of the test facility for sound insulation measurement.

II. FORMULATION

A. Representation of acoustic and vibration fields

Figure 2 illustrates the two-dimensional acoustic fields arising from an incident plane wave with angle θ. A single thin panel of finite height, h, and infinite width is placed in a rigid tunnel with a simply supported edge condition. Here, the harmonic term, $\exp(-i\omega t)$, will be omitted for convenience throughout. In region (I) depicted in Fig. 2, the total acoustic field then consists of the incident wave p_I^i, reflected wave p_I^r when the aperture is completely closed, and scattered wave p_I^s generated by the field in the aperture, which are respectively written as

$$p_I^i(x,y)=e^{ik\cos\theta(x+l_1)+ik\sin\theta y}, \tag{1}$$

$$p_I^r(x,y)=e^{-ik\cos\theta(x+l_1)+ik\sin\theta y}, \tag{2}$$

$$p_I^s(x,y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_I^s(\zeta)e^{i\zeta y-ik_{x\zeta}(x+l_1)}\,d\zeta, \tag{3}$$

where $k_{x\zeta}=\sqrt{k^2-\zeta^2}$ and k is the wave number. Note that $p_I^s(-l_1,y)$ and $\tilde{p}_I^s(\zeta)$ are the Fourier-transform pair with the following relationships:

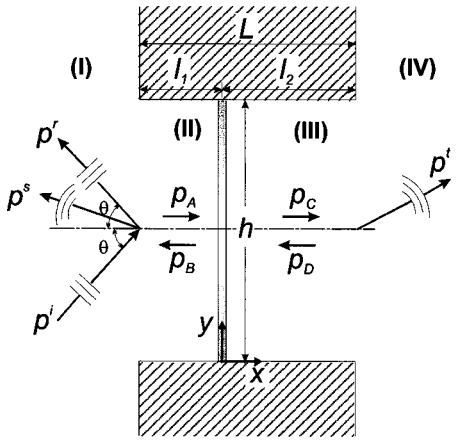


FIG. 2. Geometrical model for calculating sound transmission loss and acoustic field generated by plane wave incidence with angle θ.

$$p_I^s(-l_1,y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_I^s(\zeta)e^{i\zeta y}\,d\zeta \quad \text{and}$$

$$\tilde{p}_I^s(\zeta)=\int_{-\infty}^{\infty}p_I^s(-l_1,y)e^{-i\zeta y}\,dy. \tag{4}$$

In regions (II) and (III), the total acoustic field can be represented by a summation of normal modes:

$$p_{II}(x,y)=\sum_{n=0}^{\infty}(A_ne^{ik_{xn}x}+B_ne^{-ik_{xn}x})\cos(a_ny),$$

$$\text{when } -l_1\leqslant x\leqslant 0 \quad \text{and } 0\leqslant y\leqslant h, \tag{5}$$

$$p_{III}(x,y)=\sum_{n=0}^{\infty}(C_ne^{ik_{xn}x}+D_ne^{-ik_{xn}x})\cos(a_ny),$$

$$\text{when } 0\leqslant x\leqslant l_2 \quad \text{and } 0\leqslant y\leqslant h. \tag{6}$$

Here, $k_{xn}=\sqrt{k^2-a_n^2}$ and $a_n=n\pi/h$. In region (IV), the total acoustic field can be expressed as the transmitted field:

$$p_{IV}^t(x,y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_{IV}^t(\zeta)e^{i\zeta y+ik_{x\zeta}(x-l_2)}\,d\zeta. \tag{7}$$

Here, $p_{IV}^t(l_2,y)$ and $\tilde{p}_{IV}^t(\zeta)$ are the Fourier-transform pair which have the same relations as in Eq. (4).

The corresponding transverse displacement, $u(y)$, of a homogeneous thin plate to which the sound pressures of regions (II) and (III) are applied is governed by

$$D\frac{\partial^4u(y)}{\partial y^4}-\rho_s\omega^2u(y)=\sum_{n=0}^{\infty}(A_n+B_n-C_n-D_n)\times\cos(a_ny), \tag{8}$$

where $D=E(1-i\eta)t^3/12(1-\nu^2)$, which is the bending stiffness. Here, ρ_s, E, η, t, and ν refer to the surface density, Young's modulus, loss factor, thickness, and Poisson's ratio of the plate, respectively. Assuming the boundary condition as simply supported, the plate displacement can be expressed as

$$u(y)=\sum_{m=1}^{\infty}E_m\sin(a_my), \tag{9}$$

where $a_m=m\pi/h$ and E_m is the amplitude of the mth mode. Substituting Eq. (9) into Eq. (8), one obtains

$$\rho_s\sum_{m=1}^{\infty}(\omega_m^2-\omega^2)E_m\sin(a_my)=\sum_{n=0}^{\infty}(A_n+B_n-C_n-D_n)\times\cos(a_ny), \tag{10}$$

where $\omega_m=\sqrt{D/\rho_s}a_m^2$ and means the *in vacuo* natural frequency of the mth mode.

B. Matching boundary conditions

From the boundary conditions of pressure continuity at the Interface at $x=-I_1$ along $0<y<h$, one can obtain

$$2e^{ik\sin\theta y}+\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_I^s(\zeta)e^{i\zeta y}\,d\zeta$$

$$=\sum_{n=0}^{\infty}(A_ne^{-ik_{xn}l_1}+B_ne^{ik_{xn}l_1})\cos(a_ny). \tag{11}$$

Normal velocity continuity gives

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}(-ik_{x\zeta})\tilde{p}_I^s(\zeta)e^{i\zeta y}\,d\zeta$$

$$=\sum_{n=0}^{\infty}ik_{xn}(A_ne^{-ik_{xn}l_1}-B_ne^{ik_{xn}l_1})\cos a_ny. \tag{12}$$

In a similar fashion, the pressure and velocity continuities at $x=I_2$ along $0<y<h$ yield

$$\sum_{n=0}^{\infty}(C_ne^{ik_{xn}l_2}+D_ne^{-ik_{xn}l_2})\cos(a_ny)$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_{IV}^t(\zeta)e^{i\zeta y}\,d\zeta, \tag{13}$$

$$\sum_{n=0}^{\infty}ik_{xn}(C_ne^{ik_{xn}l_2}-D_ne^{-ik_{xn}l_2})\cos(a_ny)$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}(ik_{x\zeta})\tilde{p}_{IV}^t(\zeta)e^{i\zeta y}\,d\zeta. \tag{14}$$

The plate displacement is coupled to the acoustic pressures of the tunnel, thus the boundary conditions of both sides of the plate give

$$\sum_{m=0}^{\infty}ik_{xm}(A_m-B_m)\cos(a_my)=\rho\omega^2\sum_{r=1}^{\infty}E_r\sin(a_ry), \tag{15}$$

$$\sum_{m=0}^{\infty}ik_{xm}(C_m-D_m)\cos(a_my)=\rho\omega^2\sum_{r=1}^{\infty}E_r\sin(a_ry), \tag{16}$$

where ρ is the density of the air.

C. Solution of the coupling problem

The first step in obtaining a transmitted power is to replace the waveguide modes, i.e., A_n, B_n, C_n, D_n, with the vibration modes of the plate in Eq. (10). Multiplying Eq. (10) by $\sin(a_ry)$ and integrating the both sides with respect to y from 0 to h, one then obtains

$$\rho_s(\omega_r^2-\omega^2)E_r=\sum_{n=0}^{\infty}(P_n^{0^-}-P_n^{0^+})Y_{nr}, \tag{17}$$

where $P_n^{0^-}=A_n+B_n$, $P_n^{0^+}=C_n+D_n$, and $Y_{nr}=[1-(-1)^{r+n}]/2(r+n)\pi+[1-(-1)^{r-n}]/2(r-n)\pi$. Once the numbers of vibration and acoustic modes are taken to be M and N, respectively, the surface pressures on both sides of the plate can be derived as the following matrix form (see Appendix A):

$$\mathbf{P}^{0^-}=\mathbf{H}^{-1}\mathbf{QRE}-\mathbf{H}^{-1}\mathbf{X}, \tag{18}$$

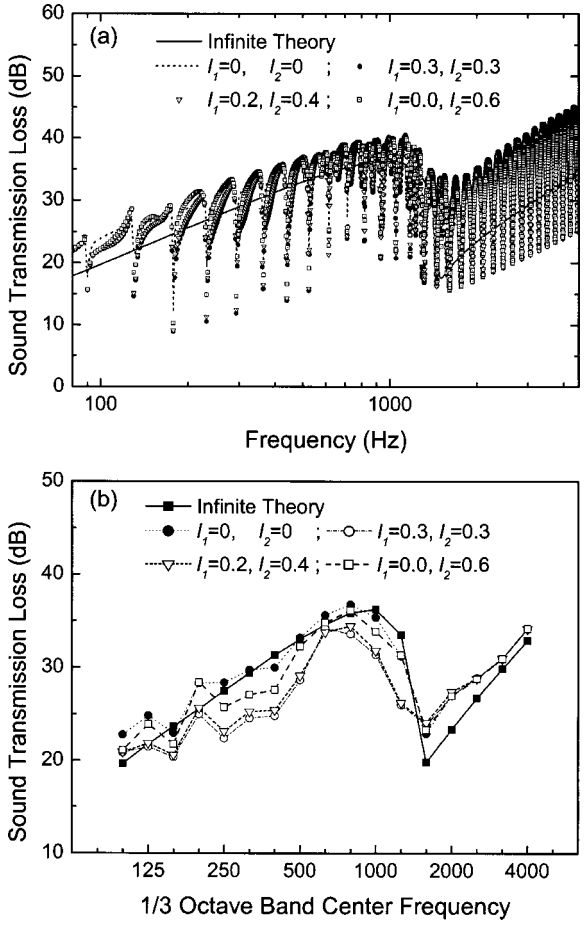


FIG. 3. Comparison of the STL between the cases without and with tunnels. (a) Narrow band and (b) $\frac{1}{3}$ octave band.

$$\mathbf{P}^{0^+}=\mathbf{S}^{-1}\mathbf{TRE}. \tag{19}$$

In the above expressions, $\mathbf{P}^{0^-}$ and $\mathbf{P}^{0^+}$ are both N-dimensional row vectors and $\mathbf{H}$, $\mathbf{Q}$, $\mathbf{S}$, and $\mathbf{T}$ are $N\times N$ matrices whose elements are defined as

$$H_{mn}=\frac{ik_{xn}J_{mn}}{2\pi}\sin(k_{xn}l_1)-\varepsilon_mh\cos(k_{xm}l_1)\delta_{mn}, \tag{20}$$

$$Q_{mn}=\frac{k_{xn}J_{mn}}{2\pi}\cos(k_{xn}l_1)-\varepsilon_mhi\sin(k_{xm}l_1)\delta_{mn}, \tag{21}$$

$$S_{mn}=\frac{ik_{xn}J_{mn}}{2\pi}\sin(k_{xn}l_2)-\varepsilon_mh\cos(k_{xm}l_2)\delta_{mn}, \tag{22}$$

$$T_{mn}=-\frac{k_{xn}J_{mn}}{2\pi}\cos(k_{xn}l_2)+\varepsilon_mhi\sin(k_{xm}l_2)\delta_{mn},$$

$$k_{xm,xn}=\sqrt{k^2-a_{m,n}^2}, \tag{23}$$

$$J_{mn}=\int_{-\infty}^{\infty}\frac{\zeta^2[1-(-1)^me^{i\zeta h}][1-(-1)^ne^{-i\zeta h}]}{(\zeta^2-a_m^2)(\zeta^2-a_n^2)}d\zeta,$$

$$m,n=0,...,N, \tag{24}$$

and $\mathbf{E}$ is a M-dimensional row vector having vibration mode coefficient, $\mathbf{R}$, $\mathbf{X}$ are $N\times M$ matrix, $N\times 1$ row vector, respectively, the elements of which can be defined by

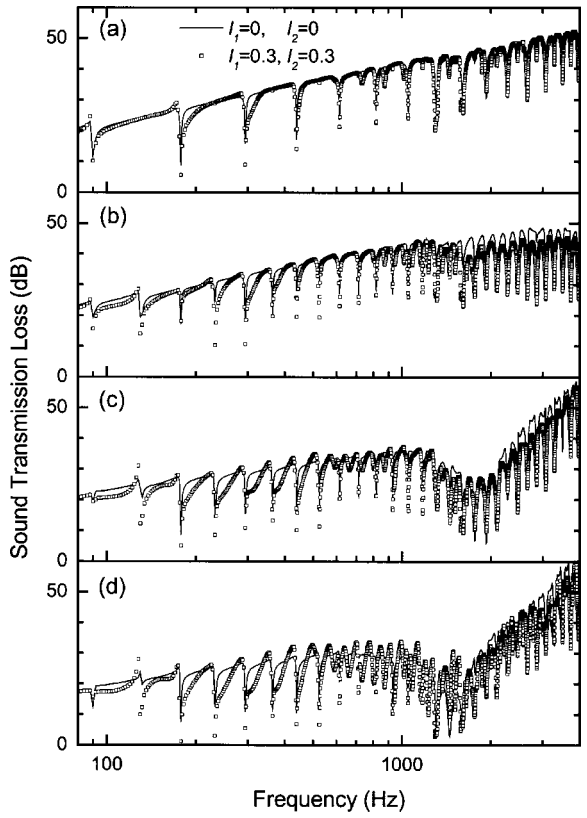


FIG. 4. Variations in STL with incidence angle. (a) θ=0°, (b) θ=30°, (c) θ=60°, and (d) θ=78°.

$$R_{nr}=\frac{\rho\omega^2}{ik_{xr}\varepsilon_r}Y_{nr}, \tag{25}$$

$$X_n=\frac{2ik\sin\theta[1-(-1)^n e^{ik\sin\theta h}]}{k^2\sin^2\theta-a_n^2}, \quad \text{where } n=0,...,N \text{ and } r=1,...,M. \tag{26}$$

Replacing Eqs. (18) and (19) into Eq. (17) allows the modal vector of plate displacement to be expressed in matrix form as

$$\mathbf{UE}=\mathbf{Y}(\mathbf{P}^{0^-}-\mathbf{P}^{0^+})=\mathbf{Y}(\mathbf{H}^{-1}\mathbf{Q}-\mathbf{S}^{-1}\mathbf{T})\mathbf{RE}-\mathbf{YH}^{-1}\mathbf{X}, \tag{27}$$

where $\mathbf{U}$ is $M\times M$ diagonal matrix with the rth element as $\rho_s(\omega_r^2-\omega^2)$. One can then obtain

$$\mathbf{E}=(\mathbf{YH}^{-1}\mathbf{QR}-\mathbf{YS}^{-1}\mathbf{TR}-\mathbf{U})^{-1}\mathbf{YH}^{-1}\mathbf{X}. \tag{28}$$

It is then possible to calculate the transmitted acoustic field from Eq. (19). The computation of Eq. (28) requires the calculation of the integral shown in Eq. (24). Unfortunately, a direct integration can not be used, since the integrand possesses singularities on the real axis. Accordingly, because the J_{mn} has four poles and two branch points in the complex ζ-plane an analytic contour integration is performed in Appendix B and gives

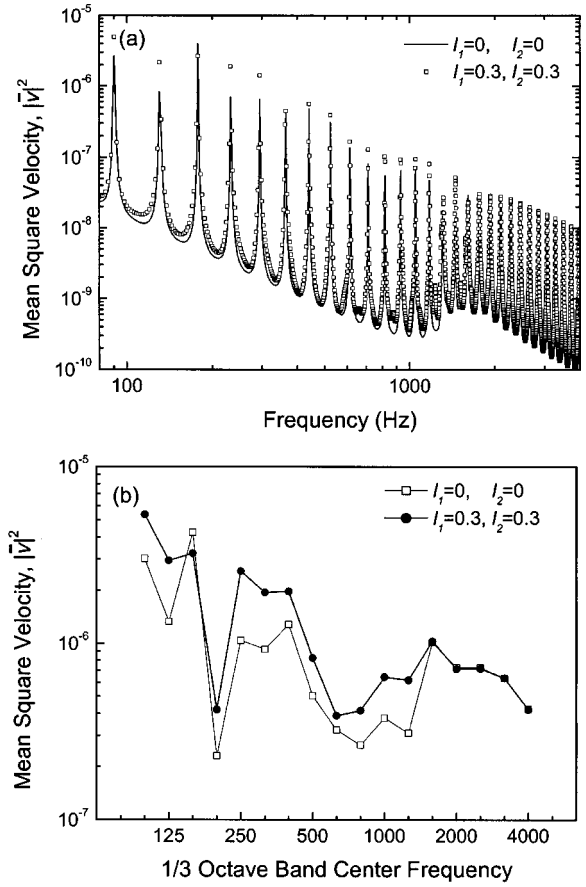


FIG. 5. Comparison of the mean square velocity between the cases without and with tunnels. (a) Narrow band and (b) $\frac{1}{3}$ octave band.

$$J_{mn}=\frac{2\pi h}{\sqrt{k^2-a_m^2}}\varepsilon_m\delta_{mn}-8i\times\int_0^\infty\frac{(1+iv^2)^2[1-(-1)^r e^{ikh}e^{-kv^2h}]}{k^2\sqrt{v^2-2i}[(1+iv^2)^2-\beta_m^2][(1+iv^2)^2-\beta_n^2]}dv, \tag{29}$$

where $\beta_{m,n}=a_{m,n}/k$. As seen from Eq. (A13), $\rho\omega J_{mn}/2\pi h\varepsilon_m$ corresponds to an acoustic modal impedance at the interface between regions (III) and (IV). Thus, the second term of Eq. (29) denotes the cross modal coupling due to the end reflection at the interface. As the frequency increases, it can be negligible compared to the first term of Eq. (29) which corresponds to the propagating modes radiating the acoustic power into region (IV) on the condition of $k^2>a_m^2$.

D. Sound transmission loss

The final step in the analysis is to find the transmission coefficient that can be expressed as the ratio of the transmitted acoustic power to the incident acoustic power. The incident acoustic power has the form of

$$\mathbf{\Pi}_i=\frac{h}{2\rho c}\cos\theta, \tag{30}$$

and the transmitted acoustic power is as follows:

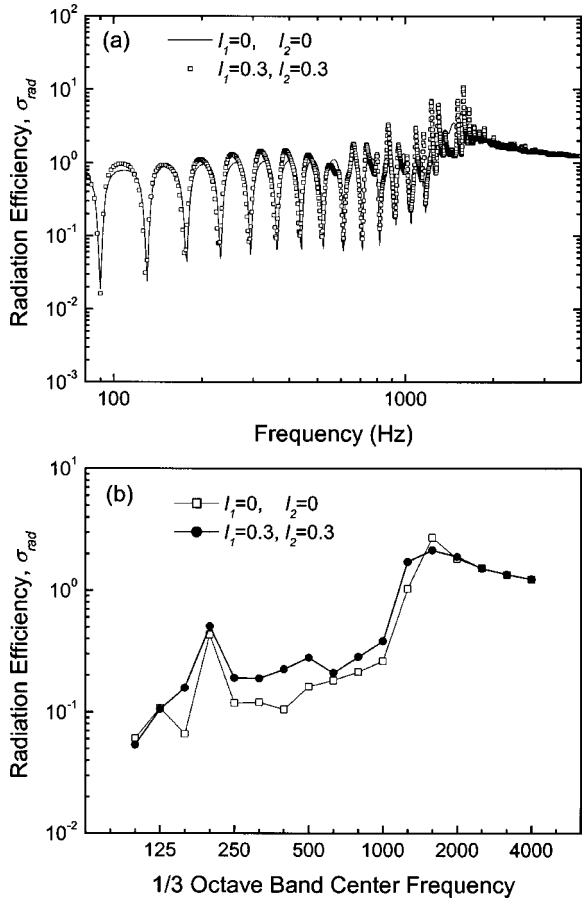


FIG. 6. Comparison of radiation efficiency between the cases without and with tunnels. (a) Narrow band and (b) $\frac{1}{3}$ octave band.

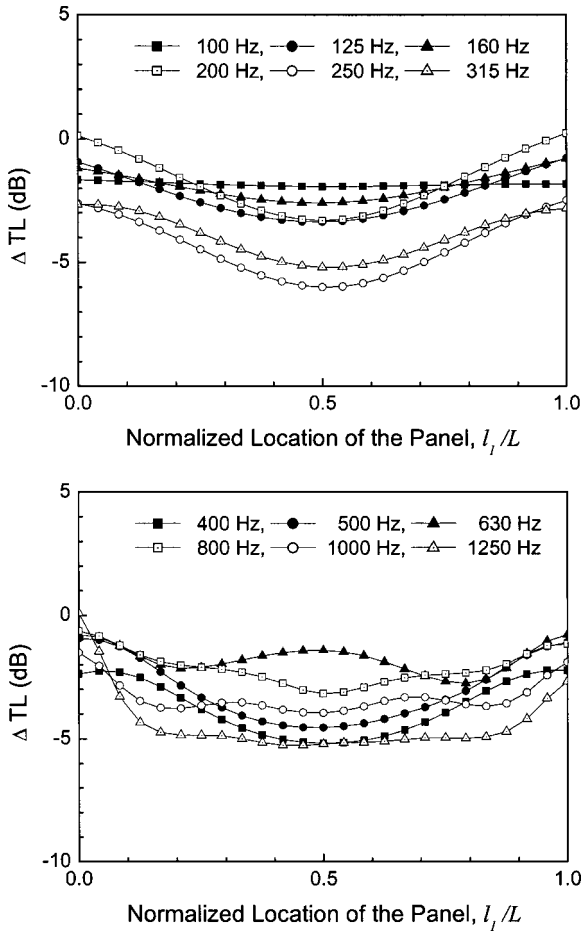


FIG. 7. ΔTL variations with the panel location in the tunnel of 0.6-m depth. Here, ΔTL is the STL difference between the cases with and without tunnels.

$$\mathbf{\Pi}_t = \frac{1}{2}\mathrm{Re}\left[\int_0^h p(l_2,y)v^*(l_2,y)dy\right]$$

$$= \frac{h}{2\rho\omega}\mathrm{Re}\left[\sum_{n=0}^{L} \varepsilon_n k_{xn}^*(C_n+D_n)(C_n-D_n)^*\right]. \quad (31)$$

Here, Re[ ] denotes the real part of [ ] and a symbol* indicates the complex conjugate. The modal coefficients C_n, D_n can be obtained from Eqs. (A13) and (A14). Since the power transmission coefficient can be given by $\tau(\theta)=\mathbf{\Pi}_t/\mathbf{\Pi}_i$, the random incidence STL can be expressed as

$$TL = 10\log_{10}(1/\tilde{\tau}),$$

$$\text{where } \tilde{\tau} = \int_0^{\theta_{\lim}} \tau(\theta)\cos\theta\, d\theta \Big/ \int_0^{\theta_{\lim}} \cos\theta\, d\theta. \quad (32)$$

Here, $\theta_{\lim}$ is the limit angle and taken as 78° throughout in the following analysis, which is based on field and laboratory measurements.[7,19]

III. NUMERICAL SIMULATIONS

In this section, the result derived earlier is applied to a glass plate with a thickness of 9.2 mm. The material properties used in the following calculation are the same as those used by Sewell,[15] i.e., $\rho t = 23\ \mathrm{kg/m^2}$, $\eta = 0.002$, and $\nu = 0.23$. To validate the formulation in Sec. II, the STL of the glass panel in the absence of a tunnel is compared to the infinite theory as shown in Fig. 3. The dips in Fig. 3(a) are due to resonance of the finite plate and frequency averaging in $\frac{1}{3}$ octave band is performed as shown in Fig. 3(b). The results in Fig. 3 indicate a reasonable agreement between infinite theory and the case without the tunnel, i.e., $(l_1,l_2)=(0,0)$, differing by less than 2 dB below the coincidence frequency.

A. Tunneling effect

Three locations of the panel were selected by varying l_1 and l_2 as defined in Fig. 1: The center position of $(l_1,l_2)=(0.3,0.3)$ m, the flush position of $(l_1,l_2)=(0,0.6)$ m, and the proportional position of $(l_1,l_2)=(0.2,0.4)$ m, which is based on the ISO 140-3 recommendation.[20] By maintaining the depth of the tunnel and the height of panel as 0.6 and 2.4 m, respectively, the STL of the plate corresponding to each location is calculated and the results are compared as depicted in Fig. 3. Below the coincidence frequency, a lower STL is obtained in all cases where the tunnel exists compared to the case without the tunnel. This is mainly due to the fact that the STL at the plate resonance dramatically decreases as shown in Fig. 3(a). In other words, the tunnel is capable of increasing resonance transmission. As pointed out by Sewell,[21] this can be explained that, unlike the case without the tunnel where a plate with an infinite baffle is excited uniformly, all the acoustic modes in the tunnel have maxima

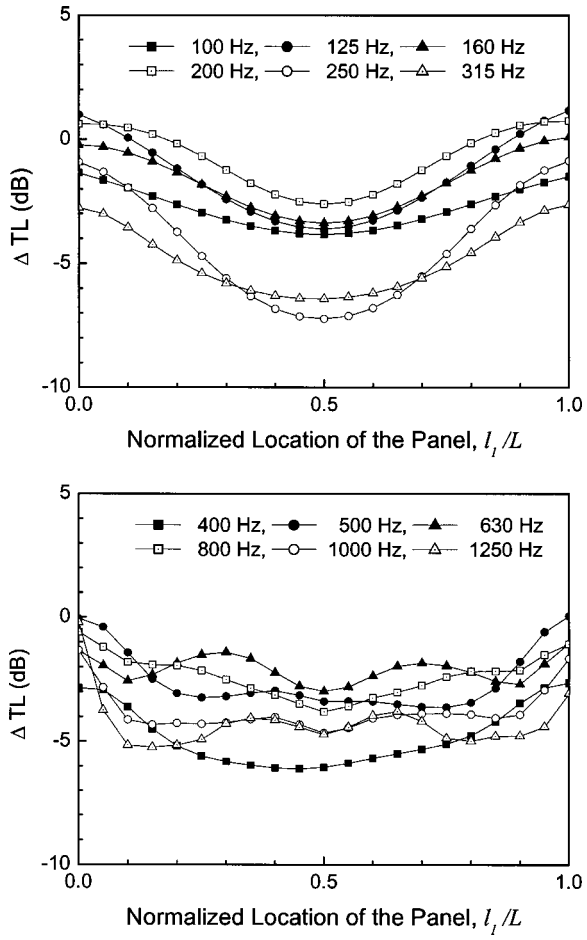


FIG. 8. ΔTL variations with the panel location in the tunnel of 1-m depth. Here, ΔTL is the STL difference between the cases with and without tunnels.

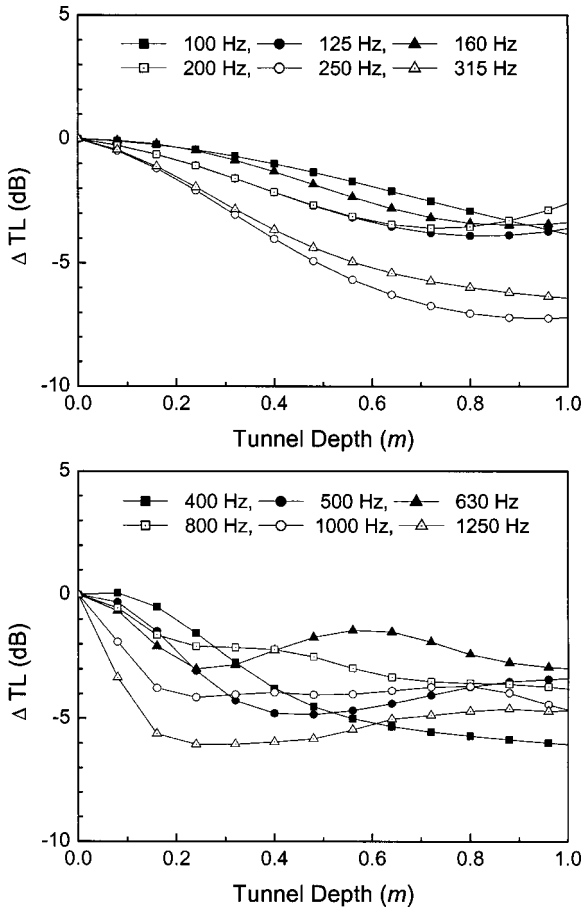


FIG. 9. Influence of tunnel depth on STL. Here, the STL change is calculated as increasing the tunnel depth where the panel is located at the center of tunnel. Here, ΔTL is the STL difference between the cases with and without tunnels.

at the edge of the panel where the resonant transmission is dominant. Another possible explanation is that end reflections from both ends of the tunnel increase the pressure difference between both sides of the plate, thus increasing the plate vibration level. Note, however, that the STL is not affected by the existence of tunnel above the coincidence. It is also noteworthy that the center-position gives the lowest STL in comparison with the others. This is due to the fact that when the geometric shapes of both tunnels are identical, the acoustic modes of each tunnel readily interact with each other through the panel, permitting the sound power to be efficiently transmitted. This explanation is generally consistent with previous results reported by Nilsson[22] and Schultz.[23]

Since the STL greatly depends on the incidence angle, the values of STL are calculated in Fig. 4 for various incidence angles as 0°, 30°, 60° and 78°. Here, a comparison is made between the case of the center-position and the case without the tunnel since the corresponding results show a large discrepancy in STL. Numerical results show that the normal incidence excites only symmetric modes of the plate, while the oblique incidence can excite both symmetric and asymmetric modes of the plate. It should be noted that the STL is greatly affected by the tunnel with increasing incidence angle. The reason for the STL decrease in the $\frac{1}{3}$ octave band is mainly due to the STL decrease at the plate resonance.

Even though this study is primarily concerned with transmission loss, a brief examination of the vibration field and radiation characteristics of the plate is carried out to examine the cause of TL change due to the presence of a tunnel. Basically, the transmitted power through the panel depends on two factors: panel vibration and radiation efficiency. The vibration fields are excited by the pressure difference across the plate as stated in Eq. (10) and the radiation efficiency can be formulated in terms of the vibration velocity of the plate and the transmitted power. The radiation efficiency, σ_{rad}, is given by

$$\sigma_{\mathrm{rad}}=\frac{\Pi_t}{\rho c h\langle\overline{v^2}\rangle}. \tag{33}$$

Here, $\langle\overline{v^2}\rangle$ represents the time and spatial average of the mean-square velocity of the plate and can be given by

$$\langle\overline{v^2}\rangle=\frac{1}{4}\sum_{n=1}^{\infty}\omega^2|E_n|^2. \tag{34}$$

By using Eqs. (33) and (34), comparisons are made between the center-position of 2.4-m height panels and the case without the tunnel since the corresponding results show a

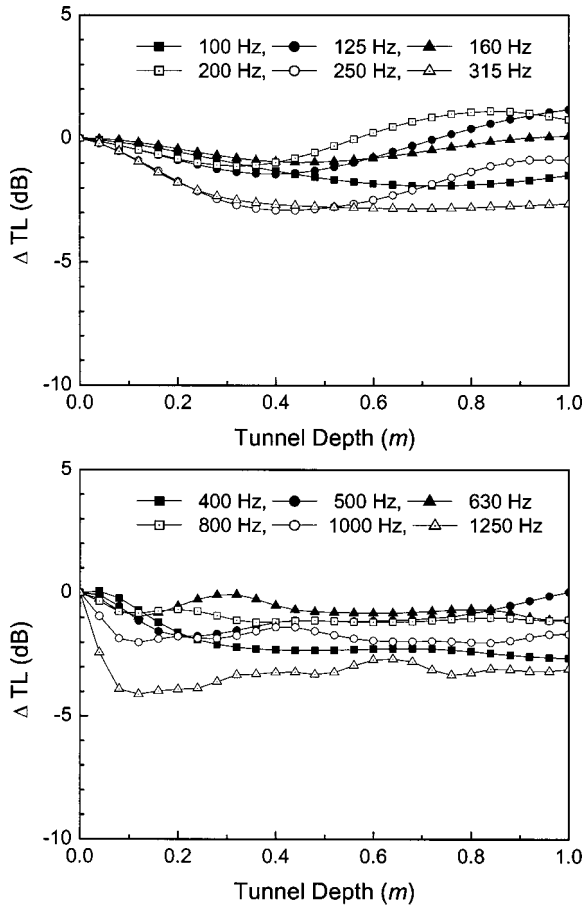


FIG. 10. Influence of the tunnel depth on STL. Here, the STL change is calculated as increasing the tunnel depth where the panel is flush with the left-end of tunnel. Here, ΔTL is the STL difference between the cases with and without tunnels.

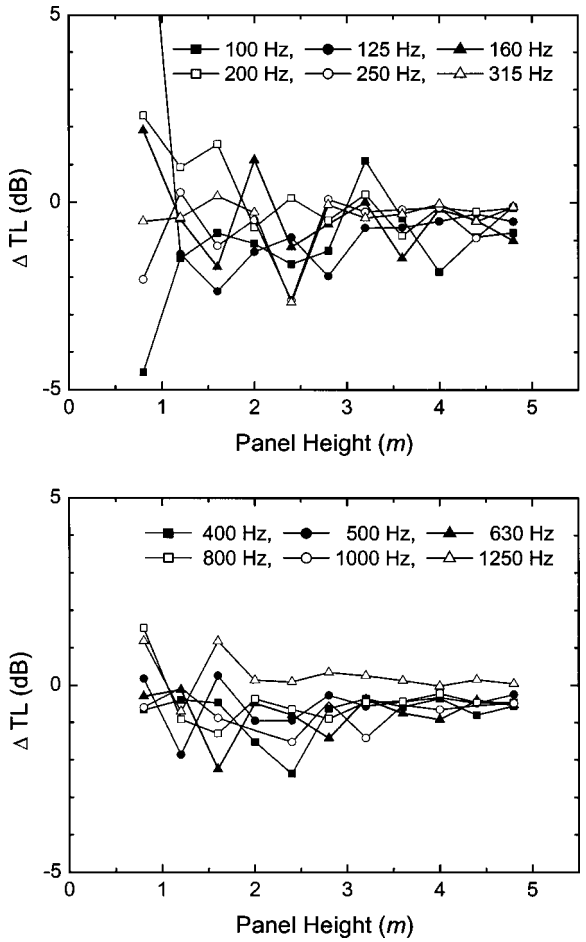


FIG. 11. STL variations with the panel size in the case that the panel is flush with the left-end of the tunnel where the tunnel depth is 0.6 m. Here, ΔTL is the STL difference between the cases with and without tunnels.

large discrepancy in STL as depicted in Fig. 3. Figures 5 and 6 show that increases in both the vibration velocity and the radiation efficiency result in a decrease in STL in the frequency range below coincidence. Above the coincidence, the values are not altered due to the presence of the tunnel and the radiation efficiency goes to unity as the frequency increases.

B. Effect of panel location

In the example presented in this section, a glass plate with a height of 2.4 m is considered throughout. For the simulation, the $\frac{1}{3}$ octave band analysis is restricted to the frequency region below 1250 Hz because the tunnel significantly alters the STL in the frequency range below coincidence. Only two different tunnel depths, i.e., 0.6 and 1 m, are used here for comparison and the changes in STL are shown as a function of normalized panel location, I_1/L. Figure 7 shows the STL difference, i.e., ΔTL, between the panels without a tunnel and with a 0.6-m depth tunnel while moving the panel from source to receiver sides. Note that the results are symmetrical with respect to the center of the tunnel. As expected, the effect of the tunnel can be minimized for the case where the panel is flush with either end of the tunnel. Additionally, by placing the panel near the center position of the tunnel, the STL is found to be remarkably decreased in the frequency range below the coincidence. A comparison between Figs. 7 and 8 also shows that the variation in STL is substantially increased with increasing tunnel depth.

C. Effect of tunnel depth and panel size

The influence of tunnel depth on the tunneling effect is investigated in detail by increasing the depth from 0 to 1 m. Figure 9 shows ΔTL of the center-located panel. The STL difference generally increases with increasing tunnel depth, especially in low frequency range. When the specimen is located at either end of the tunnel, the tunneling effect is greatly reduced compared to those of the center-located panel, although ΔTL gradually increases with tunnel depth as shown in Fig. 10. Hence, it should be noted that mounting a panel flush with the end of the tunnel can possibly improve the reproducibility of the STL obtained from different laboratories equipped with different tunnel depths.

In order to more fully comprehend the tunneling effect, a simulation of the STL with varying panel size was conducted as shown in Fig. 11. Keeping the tunnel depth fixed as 0.6 m and locating the panel in a position flush with the end of the tunnel, the height of the panel is changed from 0.8 to 4.8 m. It can be seen that noticeable STL differences are observed as the panel height becomes smaller at a frequency

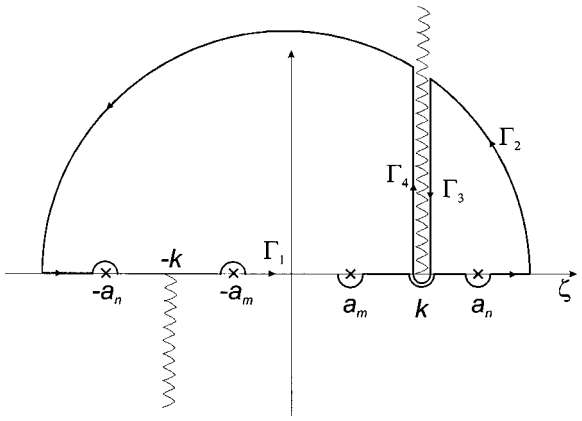


FIG. 12. Integration contour for the calculation of J_{mn}.

below 315 Hz. Again, remarkable variations in STL are observed as the frequency is decreased.

IV. CONCLUSIONS

A theoretical analysis has been made on how a tunnel affects the STL. The presence of a tunnel can exert a significant influence on the STL measurements owing to its effect on panel vibration and radiation efficiency. Basically, the tendency was to increase the vibration and radiation efficiency of the plate which, in turn, reduce the transmission loss. It is worthy to note that locating the panel flush with either end of the tunnel results in increase of the transmission loss and, in particular, gives a better solution in comparison with the analytical solution of a finite plate with the rigid baffle. As may therefore be expected, a definite increase in the STL with increasing the tunnel depth and decreasing the panel size was found. Consequently, the results confirm that the tunneling effect plays an important role in STL measurements and illustrate the possibility that the tunneling effect can be reduced, thus enhancing the reproducibility of STL measurements.

As a future work, it can be suggested that STL test of the single panel needs to be carried out in more detail to validate the present results. In addition, it certainly would appear that future work should be carried out to assess the tunneling effect in the situation where the STL is more sensitive to the incident angle, for example, a double-leaf panel with air cavity.

ACKNOWLEDGMENT

This work was partially supported by a grant from the Critical Technology Project of the Ministry of Science and Technology, Korea (M1-0139-00-0006).

APPENDIX A: CALCULATION OF SURFACE PRESSURES ON BOTH SIDES OF THE PLATE

Multiplying Eq. (11) by $\cos(a_m y)$ and integrating the both sides with respect to y from 0 to h, one then obtains

$$\frac{2ik\sin\theta[1-(-1)^m e^{ik\sin\theta h}]}{k^2\sin^2\theta - a_m^2} + \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_l^s(\zeta)\frac{i\zeta[1-(-1)^m e^{i\zeta h}]}{\zeta^2-a_m^2}d\zeta = \varepsilon_m h(A_m e^{-ik_{xm}l_1} + B_m e^{ik_{xm}l_1}), \tag{A1}$$

where $\varepsilon_m = 1$ when $m=0$ and $\varepsilon_m = 0.5$ when $m \neq 0$. Taking the inverse Fourier transform on the both sides of Eq. (12), one gets

$$(-ik_{x\zeta})\tilde{p}_l^s(\zeta) = \sum_{n=0}^{\infty} ik_{xn}(A_n e^{-ik_{xn}l_1} - B_n e^{ik_{xn}l_1})\frac{(-i\zeta)[1-(-1)^n e^{-i\zeta h}]}{\zeta^2-a_n^2}. \tag{A2}$$

Substituting Eq. (A2) into Eq. (A1), then

$$\frac{1}{2\pi}\sum_{n=0}^{\infty} k_{xn}(-A_n e^{-ik_{xn}l_1} + B_n e^{ik_{xn}l_1})J_{mn} = \varepsilon_m h(A_m e^{-ik_{xm}l_1} + B_r e^{ik_{xm}l_1}) - \frac{2ik\sin\theta[1-(-1)^m e^{ik\sin\theta h}]}{k^2\sin^2\theta - a_m^2}, \tag{A3}$$

where

$$J_{mn} = \int_{-\infty}^{\infty}\frac{\zeta^2[1-(-1)^m e^{i\zeta h}][1-(-1)^n e^{-i\zeta h}]}{(\zeta^2-a_m^2)(\zeta^2-a_n^2)}d\zeta. \tag{A4}$$

In a similar fashion, multiplying Eq. (15) by $\cos(a_n y)$ and integrating both sides with respect to y from 0 to h, one then obtains

$$A_n - B_n = \frac{\rho\omega^2}{ik_{xn}\varepsilon_n}\sum_{r=1}^{\infty} Y_{nr}E_r, \quad \text{where}$$

$$Y_{nr} = \frac{1-(-1)^{r+n}}{2(r+n)\pi} + \frac{1-(-1)^{r-n}}{2(r-n)\pi} \quad \text{and}$$

$$k_{xn} = \sqrt{k^2-a_n^2}. \tag{A5}$$

Substituting Eq. (A5) into Eq. (A3), finally, a matrix/vector equation can be obtained as follows:

$$\mathbf{P}^{0^-} = \mathbf{H}^{-1}\mathbf{Q}\mathbf{R}\mathbf{E} - \mathbf{H}^{-1}\mathbf{X}. \tag{A6}$$

In the above expression, $\mathbf{P}^{0^-}$ is a N-dimensional row vector and $\mathbf{H}$, $\mathbf{Q}$ are $N \times N$ matrices whose elements are defined by

$$P_n^{0^-} = A_n + B_n,$$

$$H_{mn} = \frac{ik_{xn}J_{mn}}{2\pi}\sin(k_{xn}l_1) - \varepsilon_m h\cos(k_{xm}l_1)\delta_{mn}, \tag{A7}$$

$$Q_{mn} = \frac{k_{xn}J_{mn}}{2\pi}\cos(k_{xn}l_1) - \varepsilon_m hi\sin(k_{xm}l_1)\delta_{mn},$$

$$k_{xm,xn}=\sqrt{k^2-a_{m,n}^2}, \quad m,n=0,...,N, \tag{A8}$$

and **E** is an M-dimensional row vector having vibration mode coefficient, and **R**, **X** are $N\times M$ matrix, $N\times 1$ row vector, respectively, whose elements are defined by

$$R_{nr}=\frac{\rho\omega^2}{ik_{xr}\varepsilon_r}Y_{nr}, \tag{A9}$$

$$X_n=\frac{2ik\sin\theta[1-(-1)^n e^{ik\sin\theta h}]}{k^2\sin^2\theta-a_n^2}, \quad \text{where}$$

$$n=0,...,N \quad \text{and} \quad r=1,...,M. \tag{A10}$$

Multiplying Eq. (13) by $\cos(a_m y)$ and integrating both sides with respect to y from 0 to h, one then obtains

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{p}_{IV}^{t}(\zeta)\frac{i\zeta[1-(-1)^m e^{i\zeta h}]}{\zeta^2-a_m^2}d\zeta = \varepsilon_m h(C_m e^{ik_{xm}l_2}+D_m e^{-ik_{xm}l_2}). \tag{A11}$$

Taking the inverse Fourier transform on both sides of Eq. (14), one obtains

$$ik_{x\zeta}\tilde{p}_{IV}^{t}(\zeta)=\sum_{n=0}^{\infty} ik_{xn}(C_n e^{ik_{xn}l_2}-D_n e^{-ik_{xn}l_2}) \times\frac{(-i\zeta)[1-(-1)^n e^{-i\zeta h}]}{\zeta^2-a_n^2}. \tag{A12}$$

Substituting $\tilde{p}_{IV}^{t}(\zeta)$ of Eq. (A12) into Eq. (A11), then

$$\frac{1}{2\pi}\sum_{n=0}^{\infty} k_{xn}(C_n e^{ik_{xn}l_2}-D_n e^{-ik_{xn}l_2})J_{rn} = \varepsilon_r h(C_r e^{ik_{xr}l_2}+D_r e^{-ik_{xr}l_2}). \tag{A13}$$

Multiplying Eq. (16) by $\cos(a_n y)$ and integrating both sides with respect to y from 0 to h, one then obtains

$$C_n-D_n=\frac{\rho\omega^2}{ik_{xn}\varepsilon_n}\sum_{r=1}^{\infty} Y_{nr}E_r. \tag{A14}$$

Substituting Eq. (A14) into Eq. (A13) and taking the numbers of vibration and acoustic modes as M and N, respectively, the matrix/vector equation can be obtained as follows:

$$\mathbf{P}^{0^+}=\mathbf{S}^{-1}\mathbf{TRE}. \tag{A15}$$

In the above expression, $\mathbf{P}^{0^+}$ is a N-dimensional row vector and **S**, **T** are $N\times N$ matrices whose elements are defined as

$$P_n^{0^+}=C_n+D_n,$$

$$S_{mn}=\frac{ik_{xn}J_{mn}}{2\pi}\sin(k_{xn}l_2)-\varepsilon_m h\cos(k_{xm}l_2)\delta_{mn}, \tag{A16}$$

$$T_{mn}=-\frac{k_{xn}J_{mn}}{2\pi}\cos(k_{xn}l_2) +\varepsilon_m hi\sin(k_{xm}l_2)\delta_{mn}, \quad m,n=0,...,N. \tag{A17}$$

APPENDIX B: CALCULATION OF INTEGRAL J_{mn} IN EQ. (24)

When m, n are opposite in parity, then $J_{mn}=0$. When m, n are the same in parity, considering Jordan's lemma,[24] J_{mn} can be rewritten as

$$J_{mn}=2\int_{-\infty}^{\infty}\frac{\zeta^2[1-(-1)^m e^{i\zeta h}]}{k_{x\zeta}(\zeta^2-a_m^2)(\zeta^2-a_n^2)}d\zeta. \tag{B1}$$

The denominator of the integrand is two-sheeted with square-root type branch points at $\zeta=\pm k$. Consider the complex ζ plane as shown in Fig. 12. Integrating along the deformed contour in the upper-half plane, it can be written that

$$J_{mn}=\frac{2\pi h}{\sqrt{k^2-a_m^2}}\varepsilon_m\delta_{mn}-4\int_{\Gamma_4}\frac{\zeta^2[1-(-1)^m e^{i\zeta h}]}{k_{x\zeta}(\zeta^2-a_m^2)(\zeta^2-a_n^2)}d\zeta, \tag{B2}$$

where the first term is the residue contribution at $\zeta=\pm a_m$ when $m=n$, and the second term is due to the integration along the branch cut Γ_3, Γ_4. By letting $\zeta=k(1+iv^2)$, the integral can be rewritten as

$$4\int_{\Gamma_4}\frac{\zeta^2[1-(-1)^m e^{i\zeta h}]}{k_{x\zeta}(\zeta^2-a_m^2)(\zeta^2-a_n^2)}d\zeta = 8i\int_0^{\infty}\frac{(1+iv^2)^2[1-(-1)^m e^{ikh}e^{-kv^2h}]}{k^2\sqrt{v^2-2i}[(1+iv^2)^2-\beta_m^2][(1+iv^2)^2-\beta_n^2]} \times dv, \tag{B3}$$

where $\beta_{m,n}=a_{m,n}/k$. Substituting the above equation into Eq. (B2), one can obtain

$$J_{mn}=\frac{2\pi h}{\sqrt{k^2-a_m^2}}\varepsilon_m\delta_{mn} -8i\int_0^{\infty}\frac{(1+iv^2)^2[1-(-1)^m e^{ikh}e^{-kv^2h}]}{k^2\sqrt{v^2-2i}[(1+iv^2)^2-\beta_m^2][(1+iv^2)^2-\beta_n^2]} \times dv. \tag{B4}$$

where $\varepsilon_m=1$ when $m=0$ and $\varepsilon_m=0.5$ when $m\neq 0$.

[1] J. D. Quirt, "Sound transmission through windows I. Single and double glazing," J. Acoust. Soc. Am. **72**, 834–844 (1982).

[2] H.-J. Kang, J.-G. Ih, J.-S. Kim, and H.-S. Kim, "Prediction of sound transmission loss through multilayered panels by using Gaussian distribution of directional incident energy," J. Acoust. Soc. Am. **107**, 1413–1420 (2000).

[3] J. S. Bolton, N.-M. Shiau, and Y. J. Kang, "Sound transmission through multi-panel structures lined with elastic porous materials," J. Sound Vib. **191**, 317–347 (1996).

[4] J. R. Callister, A. R. George, and G. E. Freeman, "An empirical scheme to predict the sound transmission loss of single-thickness panels," J. Sound Vib. **222**, 145–151 (1999).

[5] ISO 140-2:1991, "Acoustics-Measurement of sound insulation in buildings and building elements-Part 2: Determination, verification and application of precision data" (1991).

[6] T. Kihlman and A. C. Nilsson, "The effects of some laboratory designs and mounting conditions on reduction index measurements," J. Sound Vib. **24**, 349–364 (1972).

[7] R. E. Jones, "Intercomparisons of laboratory determinations of airborne sound transmission loss," J. Acoust. Soc. Am. **66**, 148–164 (1979).

[8] J. Lang, ''A round robin on sound insulation in buildings,'' Appl. Acoust. **52**, 225–238 (1997).
[9] ASTM E 413-87:1994, ''Classification for rating sound insulation'' (1994).
[10] B. Rasmussen, ''Repeatability and reproducibility of sound insulation measurements,'' Danish Acoustical Institute Technical report 118 (1984).
[11] S. D. Kristensen and B. Rasmussen, ''Measurement of sound reduction index for glazings in a staggered test opening,'' Danish Acoustical Institute Technical report 119 (1984).
[12] R. E. Halliwell and A. C. C. Warnock, ''Sound transmission loss: Comparison of conventional techniques with sound intensity techniques,'' J. Acoust. Soc. Am. **77**, 2094–2103 (1985).
[13] R. W. Guy and P. Sauer, ''The influence of sills and reveals on sound transmission loss,'' Appl. Acoust. **17**, 453–476 (1984).
[14] A. C. C. Warnock, ''Influence of specimen frame on sound transmission loss measurement,'' Appl. Acoust. **15**, 307–314 (1982).
[15] E. C. Sewell, ''Transmission of reverberant sound through a single-leaf partition surrounded by an infinite rigid baffle,'' J. Sound Vib. **12**, 21–32 (1970).
[16] D. Takahashi, ''Effects of panel boundness on sound transmission problems,'' J. Acoust. Soc. Am. **98**, 2598–2606 (1995).
[17] S. H. Ko, ''Modal contributions of a finite plate to power spectra,'' J. Acoust. Soc. Am. **87**, 1948–1954 (1990).
[18] L. D. Pope and R. C. Leibowitz, ''Intermodal coupling coefficients for a fluid-loaded rectangular plate,'' J. Acoust. Soc. Am. **56**, 408–415 (1974).
[19] L. L. Beranek, *Noise Reduction* (McGraw–Hill, New York, 1971), Chap. 13.
[20] ISO 140-3:1995, ''Acoustics-Measurement of sound insulation in buildings and building elements-Part 2: Laboratory measurements of airborne sound insulation of building elements'' (1995).
[21] E. C. Sewell, ''Exact solution for transmission of reverberant sound through a circular panel in a waveguide,'' J. Sound Vib. **12**, 397–409 (1970).
[22] A. C. Nilsson, ''Reduction index and boundary conditions for a wall between two rectangular rooms. Part I: Theoretical results,'' Acustica **26**, 1–18 (1972).
[23] T. J. Schultz, ''Diffusion in reverberation rooms,'' J. Sound Vib. **16**, 17–28 (1971).
[24] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Academic, New York, 1995).

Shear material property determination from underwater acoustic panel tests

Jean C. Piquette[a)]
Naval Undersea Warfare Center, Division Newport, 1176 Howell Street, Newport, Rhode Island 02841

(Received 29 August 2003; revised 20 January 2004; accepted 27 January 2004)

A technique is presented for determining, from underwater acoustic panel tests, passive-material shear sound speed and attenuation as functions of frequency and temperature. The frequency band of 10–200 kHz was used for development, but the applicability of the techniques described is not limited to that range. The method uses measured values of oblique-incidence insertion loss (and the corresponding change in phase) and can also take advantage of oblique-incidence echo-reduction data. Initial values for the longitudinal properties are determined by least-squares fitting of a causal material model (incorporated into a layer model) to normal-incidence insertion-loss and phase-change data. Shear properties are determined similarly, using a small number of oblique incidence angles. The method assumes a single-layer panel. A revised method of performing echo-reduction measurements that permits its accurate determination at oblique angles of incidence is also described. The angle of specular reflection is not equal to the rotation angle in the standard echo-reduction measurement geometry, and it is important to properly account for this. Consequently at a given angle of panel rotation it is necessary to use one value of the incidence angle for evaluating echo-reduction data, but another value of the incidence angle for evaluating insertion-loss data. [DOI: 10.1121/1.1687833]

PACS numbers: 43.58.Dj, 43.58.Vb, 43.60.Qv [AJZ] Pages: 2110–2121

I. INTRODUCTION

Recently, a method for measuring longitudinal properties of passive acoustical materials was described.[1] (The longitudinal sound speed is that controlled by the plane wave modulus.) A key aspect of the method is the measurement of the change in phase produced by inserting a test panel between a source and receiver. The phase change is quite sensitive to the material properties of the panel, and this sensitivity furnishes much of the accuracy of the technique. The approach involves the determination of the parameters of a causal material model, which is incorporated into a layer model of the panel, by least-squares fitting to the measurements. (A reference to the layer model being used here can be found in the literature citations given in Ref. 1.) Both the insertion loss and the phase change are simultaneously used in the fits. Four or fewer model parameters are least-squares adjusted to accommodate measurements across a wide frequency band (usually covering at least a decade), and often just two adjustable parameters total suffice to cover the entire band. The interested reader is directed to Ref. 1 for a more complete description of the technique.

An extension of the method of Ref. 1 is described here. This extension allows determination of shear material properties in addition to the longitudinal properties by including measurements obtained at oblique angles of incidence. A preliminary generalization of the method to subsume echo-reduction data obtained at oblique incidence angles is also considered, although obtaining echo-reduction measurements is not essential for determining shear properties. Other methods for determining shear material properties are also available,[2–6] but none of these methods cover the full frequency range considered here, nor do they incorporate the full universe of measurements accommodated by the present method.

The method for determining shear material properties is described in Sec. II, and the technique for obtaining and analyzing echo-reduction measurements at oblique incidence angles is described in Sec. III. Experimental measurements and applications are presented in Sec. IV. A discussion of the results and related work is given in Sec. V, and Sec. VI gives a summary and the conclusion.

II. METHOD FOR DETERMINING SHEAR MATERIAL PROPERTIES

The experimental arrangement is essentially the one typically used to measure insertion loss,[7] although an additional ''reference'' hydrophone is also required for accurate phase measurement.[1] As with the method described in Ref. 1, measurements must be acquired both with the panel present and with the panel absent. Such measurements are required to determine both the phase change and the insertion loss. Here, the term ''phase change'' refers to the phase angle that would be associated with the ratio formed in computing insertion loss, if complex amplitudes were used. In the conventional definition of insertion loss, which is typically a real quantity, a ratio is formed between the wave amplitude with the panel absent to the wave amplitude with the panel present. Twenty times the logarithm to the base 10 of this ratio is the usual definition of insertion loss.[7] If complex values are used, the phase angle associated with the ratio is the change in phase induced by the insertion of the panel between source and receiver. As defined here, the phase

a)Electronic mail: piquettejc@npt.nuwc.navy.mil.

change is taken to be the measured phase of the pressure field when the panel is absent minus the measured phase when the panel is present.

It is assumed that the insertion loss and phase change initially have been measured at normal incidence, and that the longitudinal properties of the sample have been determined, using the method of Ref. 1, for a wide range of test frequencies. Next the insertion loss and phase change are measured at a small number of oblique incidence angles, using the same test frequencies at which the normal-incidence measurements were made. It is assumed that shear-wave generation is sufficient at the oblique incidence angles to produce significant differences from the results obtained at normal incidence. A causal material model, analogous to that introduced for the longitudinal phase speed and attenuation, is introduced to represent shear material behavior. Specifically, the (unknown) shear attenuation function is represented by a Maclaurin series expansion truncated at the quadratic term. The associated phase speed is computed analytically from the attenuation function by applying suitable Kramers–Kronig relationships.[8] The resulting expressions relating phase speed and attenuation are the same as those obtained for the longitudinal phase speed and attenuation in Ref. 1. These equations are reproduced here for the reader's convenience. The equation for the shear attenuation is

$$\alpha^{\text{shear}}(\omega)=\frac{a_0^{\text{shear}}+a_1^{\text{shear}}(\omega/\omega_0)+a_2^{\text{shear}}(\omega/\omega_0)^2}{[1+b(\omega/\omega_0)^4]}, \tag{1}$$

where a_0^{shear}, a_1^{shear}, and a_2^{shear} are the unknown coefficients in the Maclaurin representation, and ω is the angular test frequency. The Kramers–Kronig relationships produce a connection between phase speed and attenuation as given by Eq. (2),

$$\frac{\omega}{c_{\text{phase}}^{\text{shear}}(\omega)}=\frac{\omega}{c_\infty^{\text{shear}}}+[\alpha^{\text{shear}}(\omega)], \tag{2}$$

where the [ ] notation denotes the Hilbert transform and c_∞^{shear} denotes the infinite-frequency shear phase speed. Substituting Eq. (1) into Eq. (2) and carrying out the Hilbert transform produces the result

$$c_{\text{phase}}^{\text{shear}}(\omega)=\frac{2b^{1/4}c_\infty^{\text{shear}}\pi(b\omega^4+\omega_0^4)}{2b^{5/4}\pi\omega^4-c_\infty^{\text{shear}}\pi\omega^2\omega_0\sqrt{b}[\sqrt{2}(a_2^{\text{shear}}+a_0^{\text{shear}}\sqrt{b})+a_1^{\text{shear}}b^{1/4}]+\omega_0^3\left\{\pi c_\infty[\sqrt{2}(a_2^{\text{shear}}-a_0^{\text{shear}}\sqrt{b})+2b^{1/4}\omega_0]-a_1^{\text{shear}}b^{1/4}c_\infty^{\text{shear}}\ln\left(b\frac{\omega^4}{\omega_0^4}\right)\right\}}. \tag{3}$$

In Eqs. (1) and (3) the parameters b and ω_0 are numerical low-pass filter characteristics that permit convergence of the Hilbert transform of Eq. (2), where the filter cutoff is set well outside the frequency interval of measurement interest. More details are furnished in Ref. 1.

Although Eqs. (1)–(3) have the same *form* as for the longitudinal case, it should be understood that the model parameters c_∞, a_0, a_1, and a_2 of the longitudinal model, and c_∞^{shear}, a_0^{shear}, a_1^{shear}, and a_2^{shear} of the shear model, have independent numerical values.

Approximate initial numerical values for the unknown shear material properties are determined in the following way: First, the longitudinal properties that have been determined at normal incidence are substituted into the layer model of the panel for numerical evaluation. This numerical evaluation is done at one of the *oblique* test angles where insertion-loss and phase-change data have been acquired. During this numerical evaluation, a two-parameter numerical search is performed for model parameters that best fit the measured phase-change and insertion-loss measurements. Since experience has shown that many acoustical materials of interest exhibit a linear variation of attenuation with frequency over a wide band, only the coefficient of the linear term a_1^{shear} in the Maclaurin-series representation of the attenuation is varied in this initial numerical search. (During this initial search the parameters a_0^{shear} and a_2^{shear} are held fixed at zero values.) Owing to the availability of Eq. (3), the only other model parameter that need be varied in the search is the infinite-frequency shear speed c_∞^{shear}. Thus, during the search only the two model parameters a_1^{shear} and c_∞^{shear} are varied.

Next, a nonlinear least-squares adjustment of the shear parameters resulting from the numerical search procedure is carried out. During this initial least-squares fitting calculation, the longitudinal properties are held fixed at the values determined from the normal-incidence data. If required to produce a satisfactory fit of the data, the model parameters a_0^{shear} and a_2^{shear} are also least-squares adjusted at this point. Finally, the longitudinal properties and shear properties are both varied simultaneously in a final least-squares fit of the measurements. If oblique-incidence echo-reduction data are available, the model is simultaneously fitted to those data as well.

In summary, the procedure is as follows:

(i) A nonlinear least-squares fit of normal-incidence insertion-loss and phase-change data determines initial values for the longitudinal material properties.

(ii) A numerical search, in which the model parameters a_1^{shear} and c_∞^{shear} are varied, but a_0^{shear} and a_2^{shear} are held fixed at zero values, is carried out to determine initial shear properties using the phase-change and insertion-loss data at one of the oblique test angles at which such data have been acquired.

(iii) A nonlinear least-squares improvement of the model properties obtained in step (ii) is done, holding the longitudinal properties fixed, using the data at the

same angle at which the numerical search was conducted. If needed, the model parameters a_0^{shear} and a_2^{shear} are also adjusted.

(iv) A final nonlinear least-squares fit, in which the longitudinal and shear properties are simultaneously adjusted, is performed. If echo-reduction data have been acquired, those data are also used in the final fit.

The material properties determined in this way are then held fixed, and the layer model is used to ''predict'' results at all angles for which measurements have been obtained. Only the incidence angle in the layer model is changed to accommodate each measurement, with no further least-squares adjustments of model properties being done. Agreement of the computed insertion loss and phase change (and, if available, echo reduction) with the full universe of measured values at all test angles where measurements were performed is taken to be a good indicator of the reliability of the properties.

III. REVISED PROCEDURE FOR MEASURING ECHO REDUCTION AT OBLIQUE INCIDENCE

In order to increase confidence in the material properties determined from the present method, it is desirable to demonstrate consistency of the properties with a large number of disparate measurements. In addition to demonstrating that a single set of material properties is consistent with the insertion-loss and phase-change measurements obtained at several angles of incidence as already described, it is also possible to demonstrate consistency of the same properties with echo-reduction measurements.[7] However, echo reduction is typically measured only at normal incidence. One reason for this is the fact that directional acoustic sources are generally used in echo-reduction measurements. If the sample is rotated, the hydrophone for echo-reduction measurements, which is usually rigged to rotate with the panel, would then be carried outside the primary beam of the source. Moreover, rotation of the panel causes the point on the panel surface from which the specularly reflected wave originates to be insonified by a different part of the source beam pattern. This is an effect that would be difficult to correct, although such correction would be especially important if the source has significant directionality. This is why an omnidirectional source is used for echo-reduction measurements in the present method. (An omnidirectional source is also used during acquisition of the required insertion-loss and phase-change data. In the case of phase-change measurements, however, the purpose of the omnidirectional source is to permit insonification of the reference hydophone.[1] Thus, the same omnidirectional sound source may be used to acquire all the data required for the present shear-property measurement method.)

The incident and reflected sound fields are usually measured simultaneously in echo-reduction tests, using suitably placed receive gates. (Gates are placed to avoid panel edge diffraction and transients.) One such gate measures the incident wave, the other such gate measures the reflected wave. Since the standard echo-reduction measurement procedure involves applying two gates to a single received waveform, we will refer to it as the ''dual-gate'' method. Such simultaneous measurements are possible due to the use of a very short-duration radiated pulse and a relatively large hydrophone-to-panel offset distance, which produces a separation-in-time of the incident and reflected waves received by the hydrophone. However, if the hydrophone is rotated out of the primary beam of the source, which may be unavoidable in oblique-incidence measurements owing to the large hydrophone-to-panel offset distance typically used in echo-reduction tests, the measurement of the incident wave using the dual-gate method will be invalid. Even if the hydrophone remains within the primary beam of the source after rotation, the incident-wave amplitude ''seen'' by the hydrophone in the dual-gate method is unequal to that which is incident upon the point of specular reflection on the panel. This is caused not only by the geometric falloff of source amplitude with propagation distance, an effect which is typically corrected (using a distance ratio) even in normal-incidence echo-reduction tests, but is also caused by the influence of the source beampattern, as mentioned above.

A modified echo-reduction test procedure was devised in order to permit use of the reflected wave at oblique incidence angles. In this modified procedure the incident wave is measured in a manner similar to the way it is measured in an insertion-loss test. That is, the test panel is removed from the rigging fixture during its acquisition. After the panel is removed, the hydrophone is rotated along a semicircular arc that displaces the hydrophone 180° from its usual position in a normal-incidence echo-reduction test. The measurement geometry is shown in Fig. 1. The large circular arc indicates how the hydrophone is repositioned after removal of the panel to acquire the incident field.

The incident-field (no-panel) measurement procedure is as follows:

(1) The test panel is rigged to the rotator shaft in such a manner that its front face lies directly on the axis of rotation.
(2) The hydrophone is also rigged to the rotator shaft, so that it will rotate through the same angle as the panel. The hydrophone is positioned at an offset distance from the geometric center of the front panel surface. This offset is typically about 34 cm.
(3) The panel is then removed from the rotator shaft, leaving the hydrophone in place.
(4) The rotator/hydrophone rig is then placed into the test facility as if an echo-reduction test of the standard type were going to be performed, although no panel is present.
(5) The rotator is then turned through a 180° angle, stopped, and a sound field is acquired (frequency domain) over the desired frequency interval of measurement interest at the 180° position. Note again Fig. 1. The incident sound field is measured with the hydrophone positioned at the end of the large arc shown.

If desired, it is also possible to avoid initially rigging the panel prior to acquiring the no-panel sound field. This alternate approach is used if no precise, predetermined, hydrophone-to-panel offset distance is required. In that case the hydrophone can be rigged at a position that will only

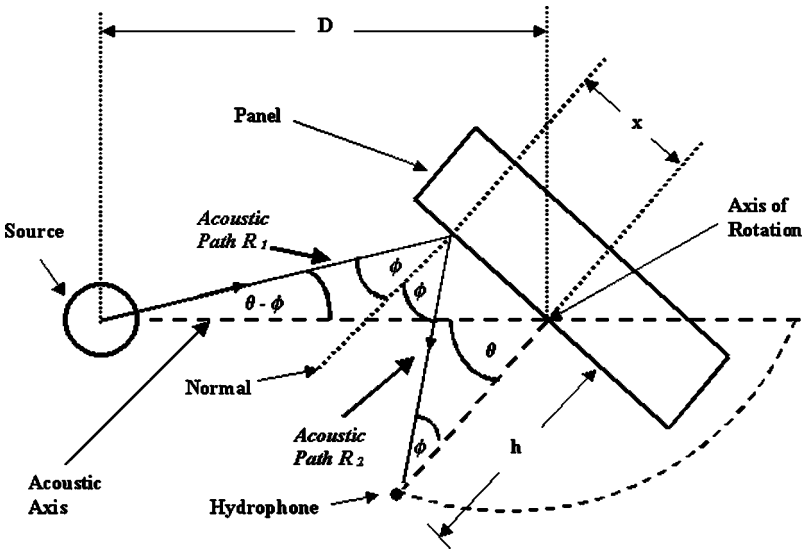


FIG. 1. Geometry of oblique-incidence echo-reduction measurements. Here, θ is the angle of panel rotation (with $\theta=0°$ denoting normal incidence); ϕ is the angle of specular reflection; D is the distance from the source to the axis of rotation; and h is the hydrophone-to-panel offset distance. The quantity x is the distance between the point on the panel surface closest to the hydrophone and the point on the panel surface from which the specularly reflected wave originates. The ''no panel'' (incident) sound field is determined by removing the panel and rotating the hydrophone to the $\theta=180°$ position. (The large dashed circular arc indicates the point to which the hydrophone is rotated, with the panel removed, to measure the incident sound field.) Distances have been exaggerated for clarity.

approximately achieve a desired hydrophone-to-panel offset after the panel is attached. If this approach is used the true hydrophone-to-panel offset distance should be measured after the panel has been rigged, since an accurately known offset distance is needed in order to allow application of the equations that follow.

Measuring the incident sound field by the revised method has three advantages over the procedure used in the dual-gate method of performing echo-reduction tests.

(i) If the hydrophone has a nonuniform beam pattern, the dual-gate method requires the use of an accurately calibrated hydrophone, even if only normal-incidence tests are of interest, in order to allow correction for the differing sensitivities in the forward and backward ''look'' directions. (The dual-gate method ''looks'' at the reflected wave on the opposite side of the hydrophone from the ''look-direction'' of the incident wave.) Since the present method involves rotating the hydrophone along a 180° arc to acquire the incident sound field, the look-direction is the same as for the reflected wave, at least in a normal-incidence test, so an uncalibrated hydrophone may be used.

(ii) In the dual-gate method, the measured incident and reflected waves travel over paths whose lengths differ substantially, owing to the extra round-trip distance that the measured reflected wave must travel between the hydrophone and the front surface of the test panel. This difference in path length usually is corrected by assuming a perfect ''$1/r$'' (spherical) falloff in source amplitude with distance. Under this assumption, which may not be valid for all source types and geometries, a distance ratio can be used to correct the data. The accuracy of this correction, though, is obviously limited by the accuracy of the $1/r$ assumption for the particular source and geometry that are being used in the test. Owing to the manner in which the hydrophone is rigged and rotated in the present method, the measured incident wave has traveled essentially the same distance through the water as the measured reflected wave, so no amplitude correction for differing path lengths must be applied. (However, see the caveat below.)

(iii) The present method permits oblique-incidence echo-reduction measurements to be performed in a manner that is similar to the way oblique-incidence insertion-loss tests are conducted, since the same no-panel sound field is used for all incidence angles.

One caveat applies to point (ii) above. This caveat derives from the fact that for oblique incidence angles the reflected wave travels along a slightly different path length than the incident wave, requiring a small geometric correction. See again Fig. 1.

The angle ϕ appearing to either side of the line labeled ''normal'' in Fig. 1 denotes the angle of specular reflection. The wave shown propagating along ''acoustic path R_1'' is the incident wave, and that propagating along ''acoustic path R_2'' is the reflected wave. The angle θ is the angle through which the panel has been rotated away from the normal-incidence position. The quantity x represents the distance between the point on the panel surface closest to the hydrophone and the point on the panel surface from which the specularly reflected wave received by the hydrophone originates. The quantity D denotes the distance of the source from the axis of rotation (which is presumed to be coincident with the front surface of the panel), and h is the offset distance of the hydrophone from front face of the panel.

Some useful equations that can be derived from the geometry of Fig. 1 are

$$\tan\phi=\frac{\sin\theta}{h/D+\cos\theta}, \tag{4}$$

$$x=h\tan\phi, \tag{5}$$

$$R_1=\sqrt{x^2+D^2-2xD\sin\theta}, \tag{6}$$

and

$$R_2=h\sec\phi. \tag{7}$$

If we let $p_{i,\mathrm{meas}}$ denote the measured amplitude of the incident wave when the hydrophone is located at the 180°, no-panel position described earlier; and if we let $p_{r,\mathrm{meas}}$ denote the measured amplitude of the reflected wave received by the hydrophone positioned as shown in Fig. 1, we can then de-

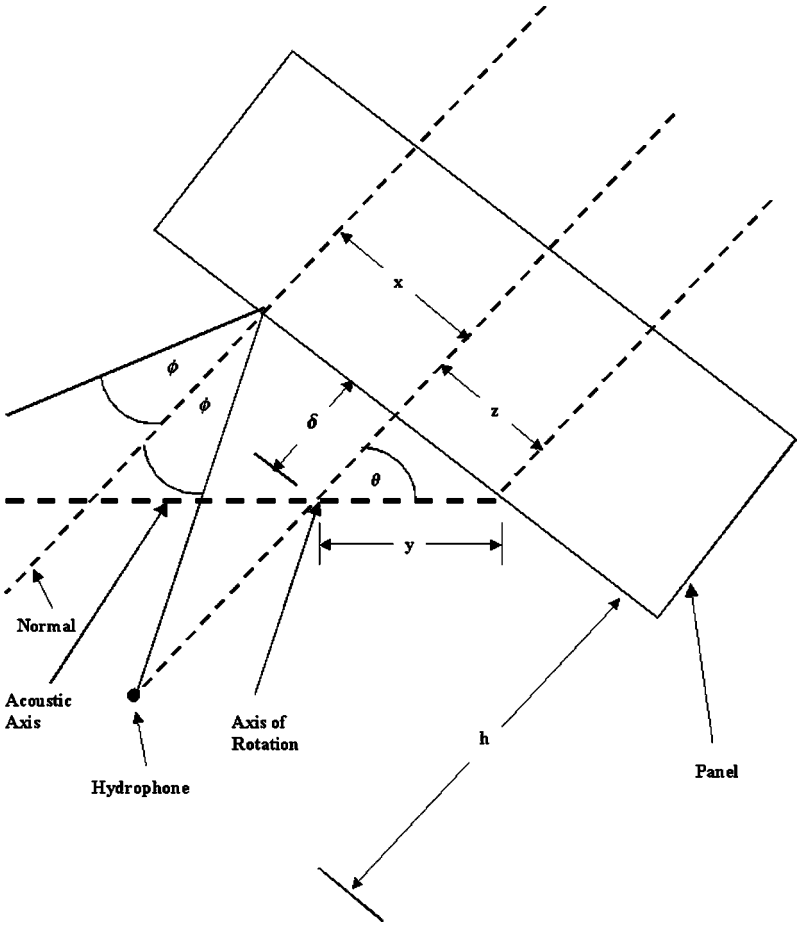


FIG. 2. Magnified view of the revised oblique-incidence echo-reduction geometry required when the front panel surface cannot be rigged coincident with the axis of rotation, but must be rigged at an offset distance δ from it.

termine magnitude of the specular reflection coefficient $|R_{\text{specular}}|$ from the measurements. The needed equation is

$$|R_{\text{specular}}| = \left(\frac{R_1+R_2}{D+h}\right) \times \left(\frac{p_{r,\text{meas}}}{p_{i,\text{meas}}}\right), \tag{8}$$

where R_1 and R_2 are given by Eqs. (6) and (7), respectively. The echo reduction follows from this equation, as the definition of echo reduction is taken in this work to be $-20\log_{10}|R_{\text{specular}}|$.

Since the two sums appearing in Eq. (8), viz. R_1+R_2 and $D+h$, are not identical, the specular reflection coefficient $|R_{\text{specular}}|$ cannot be computed by a direct ratio of the amplitudes of the measured incident and reflected fields. But the correction of this ratio given by Eq. (8) is quite small, being less than 0.15 dB for the typical values of $D=200$ cm, $h=34$ cm, and $\theta=30°$; and is less than 0.33 dB for these same distances and $\theta=45°$.

If the panel cannot be rigged so its front surface is coincident with the axis of rotation as shown in Fig. 1, but must be rigged so that its front surface is located at an offset distance δ from the axis, Eqs. (4), (6) and (8) must be revised. The revised equations are, respectively,

$$\tan\phi = \frac{\sin\theta - \dfrac{z}{D+y}}{\dfrac{h}{D+y}+\cos\theta}, \tag{9}$$

$$R_1 = \sqrt{(x+z)^2+(D+y)^2-2(x+z)(D+y)\sin\theta}, \tag{10}$$

and

$$|R_{\text{specular}}| = \left(\frac{R_1+R_2}{D+h-\delta}\right) \times \left(\frac{p_{r,\text{meas}}}{p_{i,\text{meas}}}\right), \tag{11}$$

where $y=\delta\sec\theta$ and $z=\delta\tan\theta$. See Fig. 2. Here, the rotation-axis-to-front-panel-surface offset distance δ is taken positive if the offset represents an *increased* distance from the source to the front panel surface, but negative if the offset represents a *decreased* distance. Also, D is still taken to represent the distance of the source from the axis of rotation, and h is still taken to represent the offset distance of the hydrophone from the front panel surface.

As can be seen from Fig. 1 and Eqs. (4) and (5), the point of specular reflection approaches the panel edge as the panel is rotated. Thus, the angle of rotation used in an oblique-incidence echo-reduction test cannot be made too large, or no specularly reflected wave will be received by the hydrophone. At a rotation angle of 45°, for a typical panel of 76 cm (30 in.) width, the specularly reflected wave originates from a point located only about 10.7 cm from the panel edge. This distance is equivalent to just one wavelength in room-temperature water at a frequency of about 14 kHz. Thus, a 45° rotation angle is perhaps the practical limit for echo-reduction tests for this particular geometry and panel size. For a panel of 76 cm width, the specularly reflected wave arises *directly from the panel edge* at a rotation angle of about 55.5°, so echo-reduction measurements must obvi-

ously be kept below this rotation angle for panels of this size, regardless of the frequency of interest.

A much more significant factor than the quite minor amplitude correction given by Eq. (8) in comparing theory and experiment for the arrangement of Fig. 1 is that the angle of specular reflection ϕ is not equal to the angle of rotation θ. See Eq. (4). For $D=200$ cm, $h=34$ cm, and $\theta=30°$, for example, Eq. (4) gives $\phi=25.76°$. Thus the angle of specular reflection differs from the angle of rotation by more than 4 degrees for this rotation angle. Greater differences occur at greater angles of incidence. Failing to account for the true incidence angle would produce a significant error.

Referring once again to Fig. 1, it is seen that in oblique-incidence echo-reduction tests the hydrophone look-direction for the reflected wave is the angle ϕ. If the hydrophone used in the test has a nonuniform beam pattern, a small error is introduced into the method from the fact that this look-direction differs from that used in measuring the no-panel case. If a panel of 76 cm width and a typical H-52 hydrophone[9] are used, however, the error is less than 0.1 dB for rotation angles up to 30°, and is less than 0.5 dB for rotation angles up to 45°, for frequencies below 100 kHz. (At a rotation angle of 30°, the hydrophone look-direction at the reflected wave is about 26°, and for a rotation angle of 45° the look-direction is about 39°.) Even greater accuracy is possible if additional no-panel sound-fields are acquired at the true angles of specular reflection of measurement interest, rather than using the single no-panel sound field acquired at the 180° position depicted in Fig. 1.

IV. EXPERIMENTAL RESULTS

The methods described here were applied to a variety of test panels of differing acoustical characteristics. Results obtained from two representative samples were selected for presentation. One sample was fabricated from polymethylmethacrylate (Lucite®). This sample was chosen because the acoustical properties of Lucite have been measured previously,[4,6] and the results of the present method could therefore be compared with the earlier results. The second sample was fabricated from Noryl®, a commercially available elastomeric product of unknown acoustical properties. The Lucite sample has a thickness of 3.2 cm ($1\frac{1}{4}$ in.), and has 76 cm×76 cm (30 in.×30 in.) lateral dimensions, and was tested over the frequency decade from 10 to 100 kHz. The Noryl sample has a 2.5 cm (1 in.) thickness, 91 cm ×91 cm (3 ft×3 ft) lateral dimensions, and was tested over the frequency decade from 20 to 200 kHz. The Lucite sample was tested at normal incidence; at a 20° rotation angle; and at a 30° rotation angle. The Noryl sample was tested at normal incidence; at a 10° rotation angle; and at a 25° rotation angle. A full set of insertion-loss, phase-change, and echo-reduction data were acquired for Lucite at all test angles. For Noryl no echo-reduction data were taken, but insertion loss and phase change were measured at all test angles.

The results for Lucite are shown in Figs. 3(a)–(c), 4(a)–(c), 5(a)–(c), 6(a) and (b), 7(a) and (b). Those for Noryl are shown in Figs. 8(a) and (b), 9(a) and (b), 10(a) and (b), 11(a) and (b), and 12(a) and (b). Solid-line curves are model results, while discrete points are measured results, except for

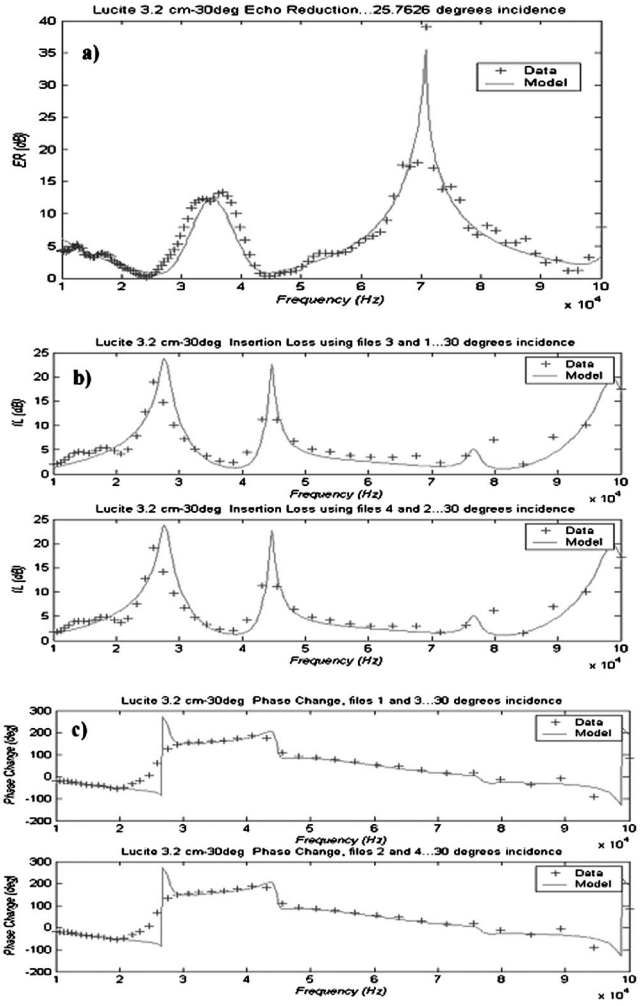


FIG. 3. Comparisons of least-squares fitted model (solid lines) to measured data (+) for the Lucite sample at the 30° rotation angle. The angle of specular reflection ϕ, as determined from Eq. (4), is actually 25.76°. It is this value of the angle ϕ, and not the 30° angle of rotation, that was used in the layer model to analyze the reflected-wave data. However, the 30° rotation angle was used for the angle of incidence to analyze the through-the-panel transmitted wave data. (a) Echo reduction. (b) Insertion loss. (c) Phase change.

Figs. 6(a) and (b), 7(a) and (b), 11(a) and (b), and 12(a) and (b) where discrete symbols denote uncertainty limits. Note that there are two plots shown for each measurement of insertion loss and phase change. These paired plots present results from two separate tests (an initial measurement and a repeat measurement), since repeating each test is essential for confirmation of measurement accuracy, owing to the difficulty of accurately performing phase-change measurements.[1] The theoretical curves that are shown in each of these paired plots (solid lines) are identical between corresponding plots, with only the measured data differing in each plot. Although insertion loss is not particularly difficult to measure, obtaining it at the same time as the phase-change measurements are repeated requires negligible additional effort, so it is also routinely recaptured when the repeat phase measurements are done.

For the Lucite sample, the least-squares fits to determine model properties were carried out using the data acquired at the 30° rotation angle. Theoretical results at the other rota-

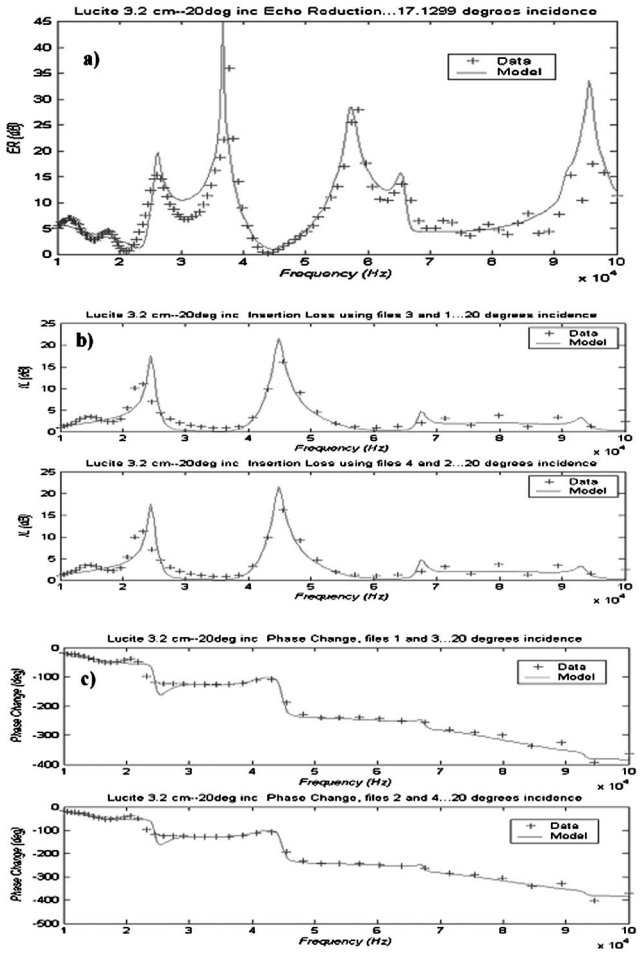


FIG. 4. Comparisons of model ''predictions'' (solid lines) to measured data (+) for the Lucite sample at the 20° rotation angle. Model predictions are based on material properties determined from the least-squares fit of the 30° rotation data, shown in Fig. 3. Only the incidence angle in the layer model was changed to accommodate the new incidence angle, i.e., no additional adjustment of model parameters was done to produce the model curves shown. The angle of specular reflection ϕ, as determined from Eq. (4), is actually 17.13°. It is this value of the angle ϕ, and not the 20° angle of rotation, that was used in the layer model in analyzing the reflected-wave data. However, the 20° rotation angle was used for the angle of incidence to analyze the through-the-panel transmitted wave data. (a) Echo reduction. (b) Insertion loss. (c) Phase change.

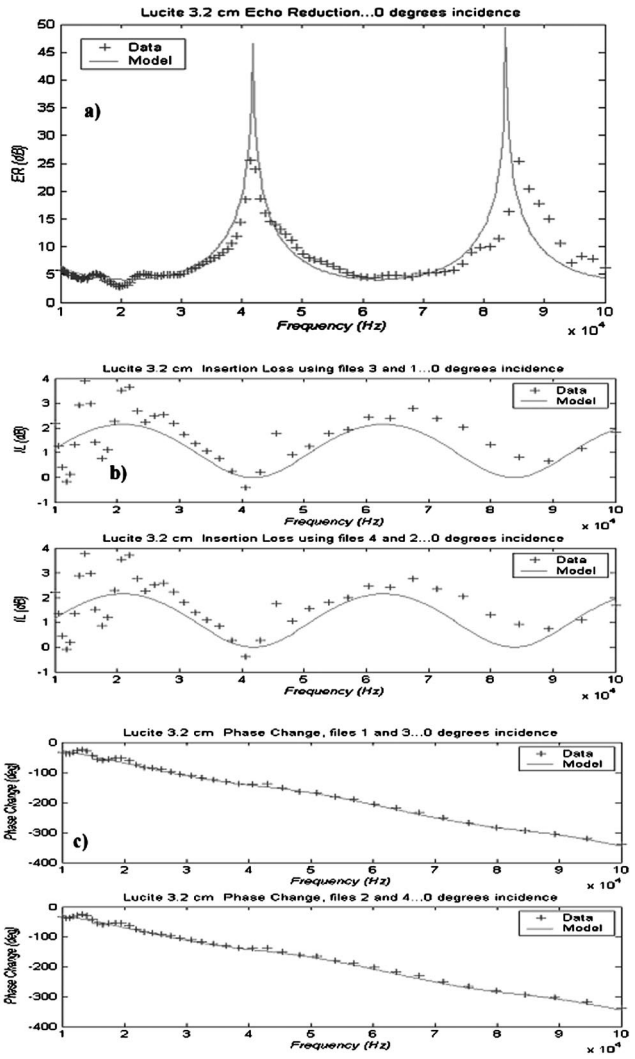


FIG. 5. Comparisons of model ''predictions'' (solid lines) to measured data (+) for the Lucite sample at normal incidence. Model predictions are based on material properties determined from the least-squares fit of the 30° rotation data. Only the incidence angle in the layer model was changed to accommodate the new incidence angle, i.e., no additional adjustment of model parameters was done to produce the model curves. (a) Echo reduction. (b) Insertion loss. (c) Phase change.

tion angles where data were acquired were computed by holding fixed the material properties determined at the 30° rotation angle. Only the theoretical incidence angle was changed in the layer model to accommodate the changed rotation angles in the remaining data sets.

A least-squares fit of the model to *all* the data available at the 30° rotation angle (echo reduction, insertion loss, and phase change at all measured frequencies) was simultaneously carried out in determining the model properties for Lucite. Only four constants total had to be adjusted in this fit (c_∞, a_1, c_∞^{shear}, and a_1^{shear}), with the remaining model parameters being held fixed at zero values. It is important to realize that the incidence angle used in the fit of the echo-reduction data [shown in Fig. 3(a)] was analytically determined from Eq. (4), rather than using the rotation angle of the measurement. Equation (4) gives $\phi = 25.76°$ for the $D = 200$ cm and $h = 34$ cm distances, and the $\theta = 30°$ rotation angle, of this test. At the same time, a 30° incidence angle (i.e., an angle equal to the angle of panel rotation) was used in the layer model for analyzing the insertion-loss and phase-change data of Figs. 3(b) and (c), respectively. The differing incidence angles for each data set are required owing to the geometry of Fig. 1, and to the differing ways in which the hydrophone is rigged for each measurement type. For the measurement of the through-the-panel transmitted wave required for evaluating insertion loss (and phase change), the hydrophone is suspended independently of the rigging that holds the panel affixed to the rotator shaft. Thus, the hydrophone that measures insertion loss (and phase change) does not participate in panel rotation. This manner of rigging the hydrophone for insertion-loss measurements is the most accurate way to measure oblique-incidence, through-the-panel transmission.[1,10] Apart from the minor influence of ''beam displacement'' [10] (i.e., the small parallel displacement of the interrogating beam resulting from refraction), when rigged in

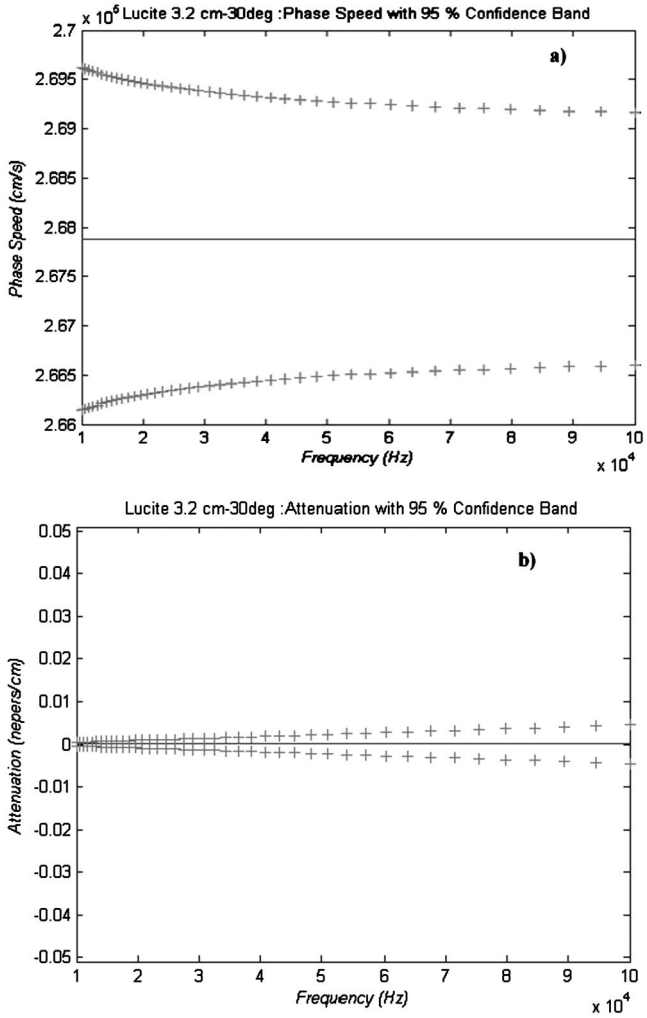


FIG. 6. Longitudinal material properties determined for the Lucite sample from a least-squares fit of the 30° rotation data. (a) Phase speed. (b) Attenuation.

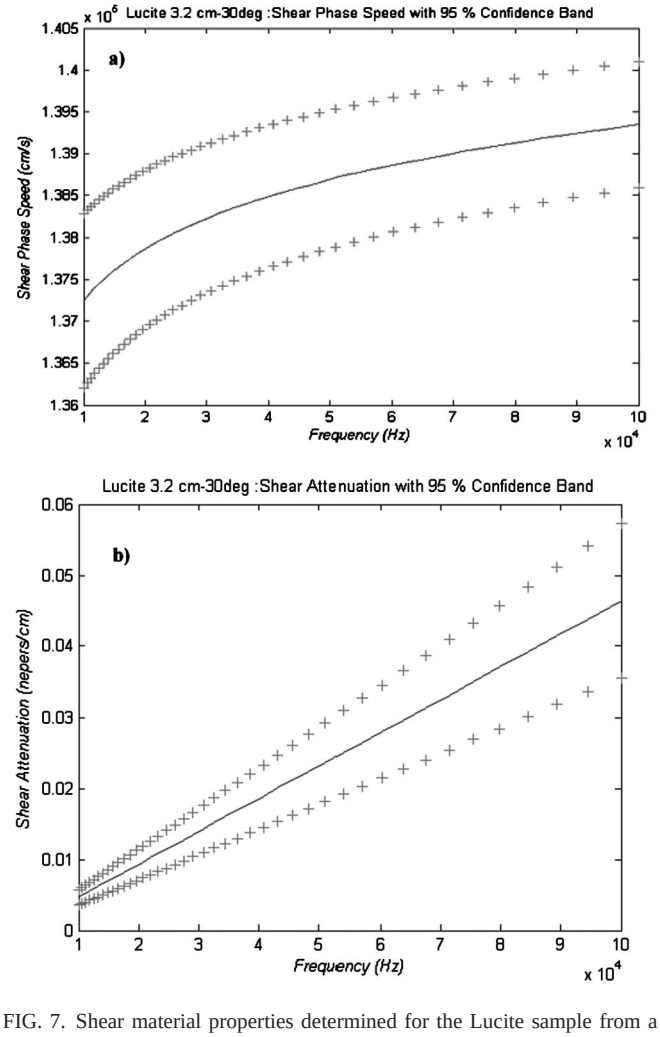


FIG. 7. Shear material properties determined for the Lucite sample from a least squares fit of the 30° rotation data. (a) Phase speed. (b) Attenuation.

this manner the through-the-panel transmitted wave may be considered to impinge upon the hydrophone in the same manner as when the panel is absent. That is, since a refraction occurs at the back panel surface that is opposite to that occurring at the front panel surface, the through-the-panel transmitted beam emerges from the panel parallel to, and only modestly displaced from, the propagation direction followed by the incident beam. It follows that the incidence angle for insertion loss is simply equal to the angle of panel rotation θ. Of course this differs from the incidence angle for specular reflection ϕ, as discussed previously. This explains why differing incidence angles are required in analyzing reflected-wave and through-the-panel transmitted-wave measurements.

To carry out the simultaneous least-squares fits of the material model to the differing data types (echo reduction, insertion loss, and phase change), a functional consisting of suitably normalized sums of the squared differences of the model and the data was formed. The functional definition is

$$f \equiv \frac{(w_{\text{phase}}/N_{\text{phase}})\Sigma_{i=1}^{n_{\text{phase}}}(\delta\phi_i^{\text{mod}}-\delta\phi_i^{\text{meas}})^2+(w_{\text{IL}}/N_{\text{IL}})\Sigma_{i=1}^{n_{\text{IL}}}(IL_i^{\text{mod}}-IL_i^{\text{meas}})^2+(w_{\text{ER}}/N_{\text{ER}})\Sigma_{i=1}^{n_{\text{ER}}}(\text{ER}_i^{\text{mod}}-\text{ER}_i^{\text{meas}})^2}{\sqrt{(w_{\text{phase}}/N_{\text{phase}})^2+(w_{\text{IL}}/N_{\text{IL}})^2+(w_{\text{ER}}/N_{\text{ER}})^2}}. \quad (12)$$

In Eq. (12) the superscript ''mod'' denotes model values and ''meas'' denotes measured values. The quantity $\delta\phi$ denotes phase change in degrees; IL denotes insertion loss; and ER denotes echo reduction, with the latter two quantities expressed in dB. The quantities denoted by N are dimensional normalization factors used to put the three disparate measurement types on an essentially equal numerical footing within the functional, and are computed by forming the sums

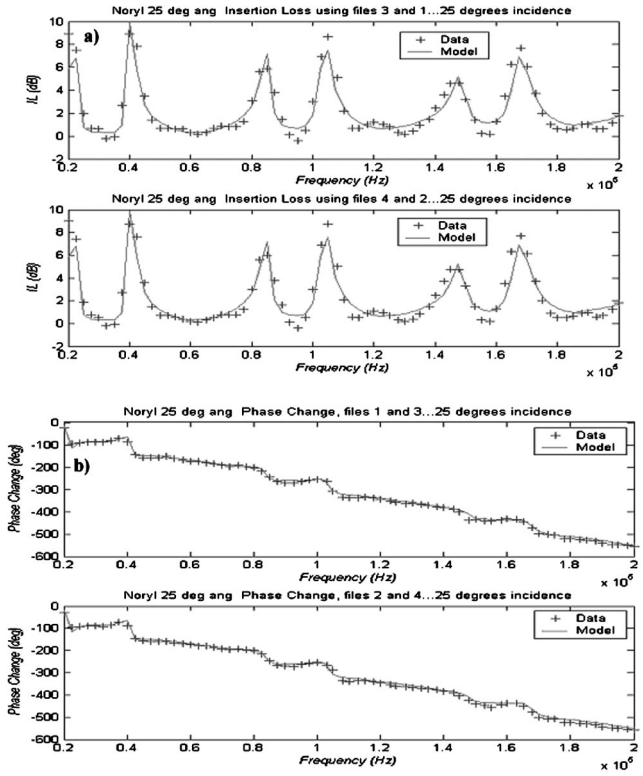


FIG. 8. Comparisons of least-squares fitted model (solid lines) to measured data (+) for the Noryl sample at the 25° rotation angle. (a) Insertion loss. (b) Phase change.

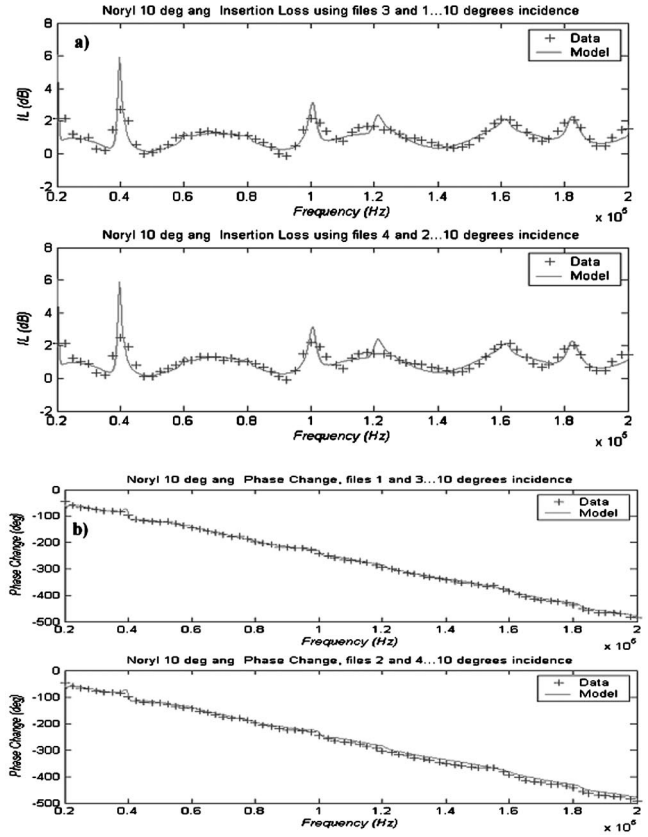


FIG. 9. Comparisons of model "predictions" (solid lines) to measured data (+) for the Noryl sample at the 10° rotation angle. Model predictions are based on material properties determined from the least-squares fit of the 25° rotation data. Only the incidence angle in the layer model was changed to accommodate the new incidence angle, i.e., no additional adjustment of model parameters was done to produce the model curves. (a) Insertion loss. (b) Phase change.

of the squares of the measured quantities indicated by the relevant subscript; n is the total number of measured values of the data type indicated by the relevant subscript; and w is a user-definable dimensionless weighting factor that can be used to increase the influence of any of the data sets on the final result. Similar procedures were used in fitting both the Lucite and Noryl samples. However, since no echo-reduction data were available for the Noryl sample, a zero weight was assigned to w_{ER} in the functional for that case, while unity values were assigned to both the phase-change and insertion-loss weights. (It should be noted that the Noryl material properties were determined using the data acquired at the 25° rotation angle.) For Lucite, an initial attempt to apply unity weighting to all three of the dimensionless weights in the functional produced an unsatisfactory fit of the model to the echo-reduction data. Specifically, the large peak of about 40 dB seen at approximately 70 kHz in Fig. 3(a) was missed by about 15 dB when using universal unity weighting, with the resulting theoretical curve passing well through the data to either side of this frequency. (That is, using unity weighting for all three data types produces an echo reduction of about 25 dB at 70 kHz, missing the peak at this frequency by about 15 dB.) To produce the fits shown, where the 70 kHz peak is missed only by about 4 dB, weights of $w_{\mathrm{phase}}=1$, $w_{\mathrm{IL}}=10$, and $w_{\mathrm{ER}}=40$ were used. The material properties deduced in the non-uniformly weighted fits differ only slightly from those obtained using unity weights for all three terms, with the most significantly affected property being the longitudinal attenuation. When unity weights are used for all terms, the resulting longitudinal attenuation is seen to increase linearly with frequency, peaking at a value of about 0.015 Np/cm at 100 kHz, rather than remaining at an essentially null value at all frequencies as seen in Fig. 6(b).

After performing the least-squares fits to the Lucite data measured at the 30° rotator angle as described, the material properties so obtained were held fixed and only the angle of incidence in the layer model was changed to generate the model curves at the other rotator angles. When the echo-reduction model curves for the 20° rotator angle were produced, as shown in Fig. 4(a), an incidence angle of $\phi=17.13°$, as given by Eq. (4) for $h=34$ cm, $D=200$ cm, and $\theta=20°$, was used in the layer-model calculations. In producing the insertion-loss and phase-change curves for this case, as shown in Figs. 4(b) and (c), respectively, the rotator angle of $\theta=20°$ was used as the incidence angle in the layer-model calculations. For the normal-incidence case, Eq. (4) produces $\phi=0$ for $\theta=0$, so a 0° incidence angle was used in the layer model in this case for producing all the curves shown in Figs. 5(a)–(c).

Some of the differences between the model and data for normal incidence [such as the apparent frequency shift between model and data seen in Fig. 5(a) in the vicinity of the peak near 85 kHz] might be attributable to spherical-wave effects, which are not accounted for in the methodology used. That is, the material properties determined from least-

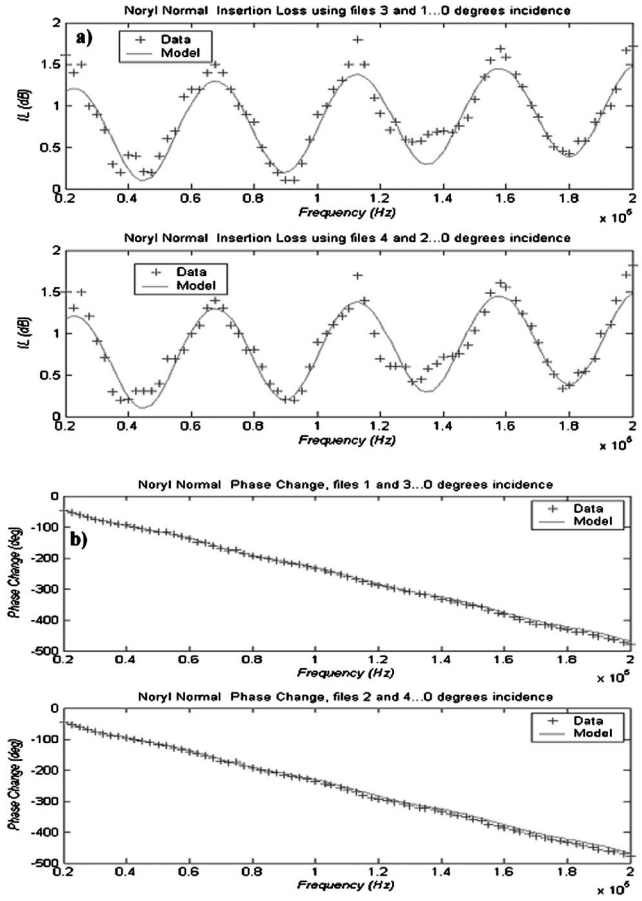


FIG. 10. Comparisons of model ''predictions'' (solid lines) to measured data (+) for the Noryl sample at normal incidence. Model predictions are based on material properties determined from the least-squares fit of the 25° rotation data. Only the incidence angle in the layer model was changed to accommodate the incidence angle, i.e., no additional adjustment of model parameters was done to produce the model curves. (a) Insertion loss. (b) Phase change.

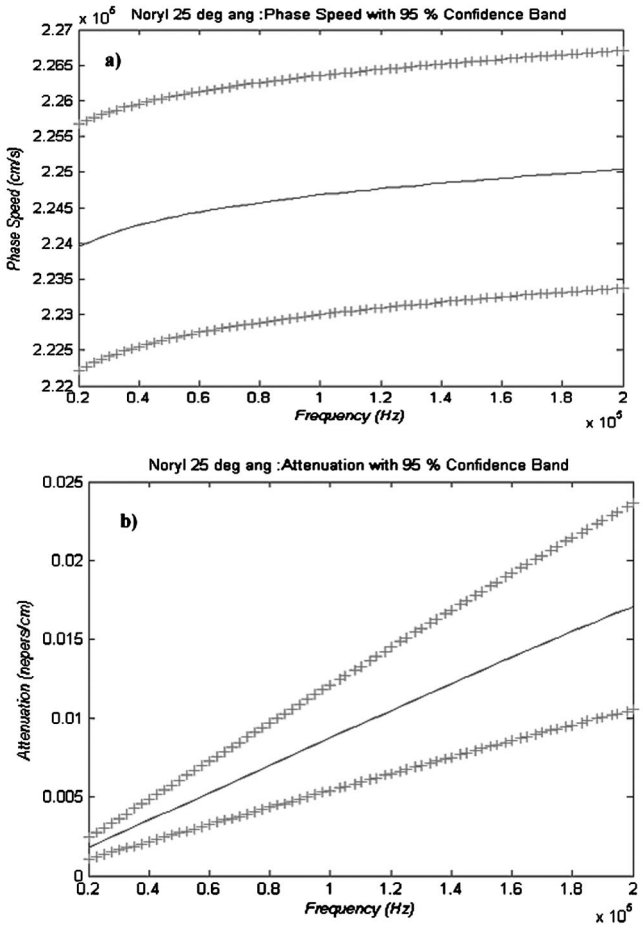


FIG. 11. Longitudinal material properties determined for the Noryl sample from a least squares fit of the 25° rotation data. (a) Phase speed. (b) Attenuation.

squares fits to the 30° test data might be affected by errors caused by spherical-wave effects at that incidence angle that differ from the spherical-wave effects present at normal incidence, possibly producing the frequency shift seen in Fig. 5(a). This issue is a matter of further investigation.

The error bars shown for the results presented in Figs. 6(a) and (b), 7(a) and (b), 11(a) and (b), 12(a) and (b) were computed using a propagation-of-error calculation based on 95% confidence limits.[11] The interested reader is directed to Refs. 1 and 11 for more details of the error calculations. It is worthwhile noting that the method described here assumes the sample thickness and density have been measured initially by standard nonacoustic methods. These quantities are not adjusted in the least-squares calculations. A numerical parameter-sensitivity study showed that, by far, the majority of the measurement error is associated with the adjusted quantities, rather than with the errors in the thickness and/or density measurements. Thus, the error bars shown here well capture the sensitivity of the computed quantities (i.e., sound speed and loss) to the uncertainties in the measurements. However, if sound speeds and losses under conditions of high hydrostatic pressure were required to be measured, it would be necessary to account for the resulting changes in sample thickness and density, either through least-squares adjustment of these quantities, or by some method of direct measurement. For example, changes in thickness could be determined performing a time-of-flight measurement on a high-frequency pulse. Changes in density could readily be estimated from changes in thickness. Of course, indirect measurements and least-squares adjustment could also be used in combination.

V. DISCUSSION

The determination of material properties from oblique-incidence panel tests has been considered previously.[2,4,12–14] In the method of Ref. 12, shear properties of the panel material were assumed negligible, except that shear-wave production generated within a support plate that may be used in the test was taken into account. The present method, however, fully accounts for shear-wave production in the material under test, although only single-layer samples can be accommodated. Moreover, the present model fully accounts for causality, through the imposition of the Kramers–Kronig relations.[1,8]

A technique similar to that considered here was proposed previously by Bar and Walti[13] and Reissner.[14] (See Mason[4] for a description of the results of these references in English.) The present method differs in several significant

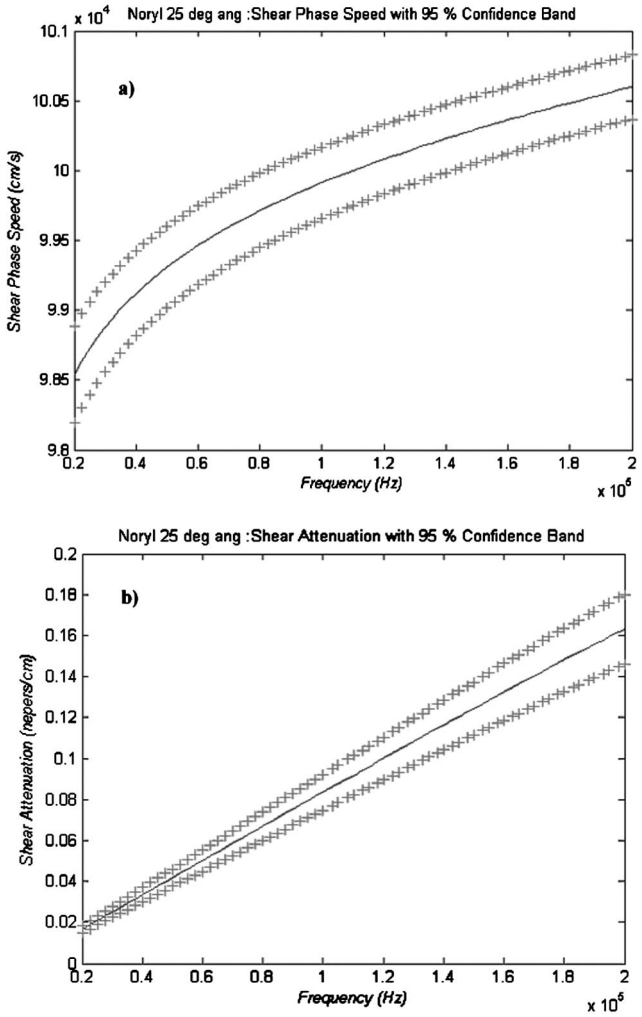


FIG. 12. Shear material properties determined for the Noryl sample from least squares fit of the 25° rotation data. (a) Phase speed. (b) Attenuation.

ways from this earlier method, and the present method has several advantages. For example, the method of Bar and Walti assumes negligible attenuation in the sample, whereas the present method allows not only for the presence of significant attenuation, but dispersion (frequency-dependent attenuation) as well. Also, the Bar and Walti method only uses measurements of the insertion loss, and only at a single frequency. The data must be obtained over a continuum of incidence angles. The present method requires measurements at only a few incidence angles, but permits accommodation of measurements taken over a continuum of frequencies. The present method also incorporates the use phase-change measurements, and the phase change is quite sensitive to the material properties of the panel. The imposition of the requirement to satisfy the causality condition in the material model used in the present method also permits straightforward accommodation of dispersive materials, i.e., materials whose properties vary significantly with frequency. Since no condition is imposed from frequency-to-frequency in the Bar and Walti method, measurement uncertainties could produce significantly acausal results in the determined properties, especially for dispersive materials where frequency variation is significant. Moreover, although the present method permits the accommodation of significant frequency variation in the material properties of the sample under test, it nonetheless requires only a very small number of adjusted parameters (four or fewer for each wave type) to cover a wide frequency band. The frequency band covered is generally a decade. The Bar and Walti method, on the other hand, requires the determination of two parameters for each of the two wave types at every frequency of the test. That is, four parameters must be determined at each frequency of the test, and a full angular sweep of the through-the-panel transmitted wave is required at each frequency. Thus, even if the Bar and Walti method were applied at the more than 100 test frequencies considered here, it would require the determination of a corresponding number of model parameters. The incorporation of echo-reduction data into the present method is also a significant improvement over the earlier technique, which relies only on insertion-loss measurements to determine the properties of the sample under test.

Although the present method and the method of Bar and Walti differ significantly, it is nonetheless useful to compare results for Lucite obtained by each technique. The Bar and Walti method was applied by Mason[4] to a Lucite sample immersed in water at a test frequency of 20 kHz. The value of the shear speed for Lucite found by Mason is 1240 m/s (see Ref. 4, p. 403), while that found in the present work is 1378 ± 11 m/s, at the same frequency of 20 kHz as considered by Mason. [See Fig. 7(a).] Of course exact agreement is not expected, since considerable variations exist between various Lucite formulations.[15,16] (See also Ref. 6, p. 158.) These differences are reflected not only in differing longitudinal and shear sound speeds and attenuations, but even in differing densities. For example, the shear sound speed deducible from the shear modulus given in Ref. 16 is about 1080 m/s, a value that is considerably less than that determined either by Mason or in the present work. On the other hand, the longitudinal sound speed of 2300 m/s given by Mason (see Ref. 4, p. 403) is considerably less than the value of 2679 ± 15 m/s found here [see Fig. 6(a)], or as found in other references.[6,14,15] For example, Refs. 6 and 15 both list the longitudinal sound speed as 2680 m/s, while Ref. 16 lists it as 2650 m/s.

While some of the differences between the results obtained by Mason and those obtained here are attributable to intrinsic differences in the Lucite samples considered, much of it may also be attributable to differences in the methodologies used. The use of phase-change measurements, and perhaps the use of echo reduction as well, may result in more accurate property determination by the present method.

It is also worthwhile considering briefly the consequences of attempting to apply the present method to high-loss materials. Clearly in order for the method to be applicable, it is necessary for sufficient acoustic energy to penetrate through the sample. If loss becomes too high, insufficient information will be available for measurement by the insertion-loss hydrophone. While it is not possible to specify initially what is an excessive level of loss, if at least approximate values of the unknown quantities are available, numerical modeling can at least provide some guidance to the design of the needed experiment. For example, for increased loss it may be necessary to reduce sample thickness.

It might also be that for a sufficiently high loss the measurement of echo reduction would become mandatory, since it is expected that the reflected wave would still be capable of measurement even if sample loss is high. In any event, an indication of the reliability of a result for a specific case would be given by the sizes of the error bars deduced by the present method. If the loss becomes too large for accurate property determination, this fact will be revealed by unacceptably large error bars.

VI. SUMMARY AND CONCLUSIONS

An initial investigation into a method for simultaneously determining longitudinal and shear sound speed and attenuation, as functions of frequency and temperature, from underwater acoustic panel tests was described. The method is based on least-squares fits of a causal material model, incorporated into a layer model, to insertion-loss, phase-change, and (possibly) echo-reduction data acquired at a small number of incidence angles. A revised technique for measuring echo reduction at oblique incidence angles was also described. This method treats the measurement of the incident sound field in an echo-reduction test in a manner that is similar to the way it is measured in an insertion-loss test. An important consideration in analyzing oblique-incidence echo-reduction data is accounting for the fact that the angle of specular reflection is not equal to the angle through which the panel has been rotated. When a layer model is applied to echo-reduction measurements acquired from a test panel, it is the computed angle of specular reflection, as given by Eq. (4), that must be used for the angle of incidence. However, the angle through which the panel has been rotated is the correct incidence angle to use for insertion-loss and phase-change data. Thus differing angles of incidence must be used in analyzing data acquired at one rotator angle, depending upon whether reflected- or through-the-panel transmitted-wave data are under consideration.

Future efforts for improving the method include the investigation of techniques for accounting for wave-front curvature, and for accommodating the influences of high hydrostatic pressure.

ACKNOWLEDGMENTS

The Office of Naval Research, Code 321 supported this work. I am grateful to Alan D. Pierce for furnishing translations of the titles of Refs. 13 and 14.

[1] J. C. Piquette and A. E. Paolero, ''Phase change measurement, and speed of sound and attenuation determination, from underwater acoustic panel tests,'' J. Acoust. Soc. Am. **113**, 1518–1524 (2003).

[2] H. J. McSkimin, ''Ultrasonic methods for measuring the mechanical properties of liquids and solids,'' *Physical Acoustics IA*, edited by W. P. Mason (Academic, New York, 1964), pp. 297–317.

[3] G. Lee, W. M. Madigosky, and J. J. Eynck, ''Dynamic viscoelastic properties of materials for underwater acoustic applications,'' Naval Surface Weapons Center report No. TR 78-138, 1979.

[4] W. P. Mason, *Piezoelectric Crystals and their Application to Ultrasonics* (Van Nostrand, New York, 1950), pp. 399–412.

[5] H. J. McSkimin and P. Andreatch, ''Measurement of dynamic shear impedance of low viscosity liquids at ultrasonic frequencies,'' J. Acoust. Soc. Am. **42**, 248–252 (1967).

[6] R. N. Capps, C. M. Thompson, and F. J. Weber, ''Handbook of sonar transducer passive materials,'' Naval Research Laboratory Memorandum Report 4311, 1981.

[7] R. J. Bobber, *Underwater Electroacoustic Measurements* (USGPO, Washington, DC, 1970), p. 287.

[8] C. W. Horton, ''Dispersion relationships in sediments and sea water,'' J. Acoust. Soc. Am. **55**, 547–549 (1974).

[9] Anonymous, *Underwater Electroacoustic Standard Transducers* (Naval Research Laboratory, Orlando, FL, 1982).

[10] J. C. Piquette, ''Conventional oblique-incidence panel test procedures,'' Naval Undersea Warfare Center Technical Memorandum 01-035, 2001.

[11] W. H. Press, A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipies in FORTRAN* (Cambridge University Press, New York, 1992), Chap. 15.

[12] J. C. Piquette, ''Offnormal incidence reflection coefficient determination for thick underwater acoustic panels using a generalized ONION method,'' J. Acoust. Soc. Am. **87**, 1416–1427 (1990).

[13] R. Bar and A. Walti, ''Uber die bestimmung der Poisson'schen elastizitatskonstante mit hilfe von ultraschallwellen'' (''Determination of Poisson's elastic constant using ultrasound waves''), Helv. Phys. Acta **7**, 658–661 (1934).

[14] H. Reissner, ''Der senkrechte und schrage durchtritt einer in einem flussigen meidum erzeugten ebenen dilatations-(longitudinal)-welle durch eine in diesem medium befindliche planparallele feste platte'' (''Perpendicular and oblique transmission of plane dilatational waves through a solid plate immersed in a fluid''), Helv. Phys. Acta **17**, 140–155 (1938).

[15] J. R. Frederick, *Ultrasonic Engineering* (Wiley, New York, 1965), p. 363.

[16] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundamentals Acoustics* (Wiley, New York, 1982), p. 461.

The eigencomponent association method for adaptive interference suppression

Brian F. Harrison[a)]
Naval Undersea Warfare Center, Submarine Sonar Department, Code 2121, Building 1320, Newport, Rhode Island 02841-0001

(Received 29 August 2003; revised 31 December 2003; accepted 19 February 2004)

The performance of passive localization algorithms can become severely degraded when the target of interest is in the presence of interferers. In this paper, the eigencomponent association (ECA) method of adaptive interference suppression is presented for signals received on horizontal arrays. ECA uses an eigendecomposition to decompose the cross-spectral density matrix (CSDM) of the data and then beamforms each of the eigenvectors. Using an estimate of the target's bearing, the target-to-interference power in each eigenvector at each CSDM update is computed to determine which are dominated by interference. Eigenvectors identified to contain low target-to-interference power are subtracted from the CSDM to suppress the interference. Using this approach, ECA is able to rapidly adapt to the hierarchical swapping of target and interference-related eigenvectors due to relative signal power fluctuations and target dynamics. Simulated data examples consisting of a target and two interferers are presented to demonstrate the effectiveness of ECA. These examples show ECA enabling accurate localization estimates in the presence of interferers, which without using the technique was not possible.
[DOI: 10.1121/1.1699395]

PACS numbers: 43.60.Fg, 43.60.Mn [EJS] Pages: 2122–2128

I. INTRODUCTION

The performance of passive localization algorithms can become severely degraded when the target of interest (TOI) is in the presence of interferers. These interferers can mask the location of the TOI when their received energy is greater than that of the TOI. This occurs because the interference energy excites ambiguous sidelobes in the localization search space, resulting in false peaks and erroneous localization estimates. In this work, we are concerned with the problem of suppressing the effects of an interferer which is in close proximity to a TOI in bearing, thereby enabling the estimation of the range and depth of the TOI. We assume data are received on a horizontal array of sensors. Bearing space affords the greatest degree of discrimination in separating targets as compared to range and depth space. Using a conventional beamformer, the beam pattern in bearing space has monotonically decreasing sidelobe levels which provide significant attenuation of interfering targets away from the main beam-steering direction. The sidelobe structure in range and depth space is more irregular in nature and provides little discrimination in resolving multiple targets. Therefore, if two targets are sufficiently separated in bearing space, their positions in range–depth space can be readily estimated even if they overlap in range or depth. As the two targets move close to each other in bearing space, the beamformer output when steered towards the TOI becomes more of a mixture of the two signals which results in degraded localization estimates. This is particularly true when the interfering target's signal is significantly stronger at the array than the TOI's.

The principal component inverse (PCI) method[1,2] has been demonstrated as a technique for adaptively nulling low-rank interference. The PCI method assumes that the interference is strong as compared to the TOI, and will therefore be isolated to the principal components of the data matrix. Thus, the interference and TOI subspaces are disjoint to a large degree, i.e., the interference and TOI signals are predominately contained in different basis vectors of the data matrix. Given that, PCI estimates the interference signal using a subspace decomposition of the data matrix and then subtracts the estimate from the data. The residual data then consist mainly of the TOI's signal and can be used in subsequent signal-processing operations such as detection and localization. Since PCI estimates the interference using the received data as opposed to attempting to model the interference, it is immune from model mismatch issues which can reduce the level of interference suppression. The most difficult step in applying PCI is estimating the rank of the interference subspace from the received data. The assumption that the interference is contained in the principal components of the data matrix may not always be valid. Relative motion of the targets and fluctuating signal levels due to changing propagation conditions can result in hierarchical swapping and mixing of the interference and TOI-related components as a function of time. A means of adaptively determining which components of the data matrix predominantly contain the interference signal is necessary for effective application of this technique. Once identified, these components can be removed from the data to effectively enhance the TOI's signal.

In this paper, we propose a technique for data received on a horizontal array which adaptively identifies the dominant interference-related components of the data matrix and removes them. This technique, which we call the eigencomponent association (ECA) method, builds on the PCI method

a)Electronic mail: harrison_bf@ieee.org

of interference suppression. The ECA method first computes an eigendecomposition of the cross-spectral density matrix (CSDM) of the array data, and then beamforms each of the eigencomponents. Assuming an independent estimate of the TOI's bearing is given, for example from a tracker, the TOI power and the interference power contained in each eigencomponent can be estimated. This allows the separation and grouping of eigencomponents associated with sources arriving at the array from different directions. Those eigencomponents with low TOI-to-interference power are then subtracted from the CSDM. Thus ECA is able to adapt to the hierarchical swapping and mixing of the eigencomponents of the CSDM. A related technique which uses the beam response of the eigenvectors of the CSDM to provide robustness to self-nulling in computing adaptive beamformer weights was presented in Ref. 3. We present the development of the ECA method in the next section followed by demonstrations of the ECA method using simulated horizontal array data.

II. THE EIGENCOMPONENT ASSOCIATION METHOD

The discrete Fourier transform of time-domain data received on an N-sensor horizontal array from Q radiating point sources can be expressed as

$$\mathbf{y}(\omega)=\sum_{k=1}^{Q} s_k(\omega)\mathbf{a}_k(\omega,\phi,\theta)+\mathbf{n}(\omega), \tag{1}$$

for frequency ω, where $\mathbf{y}(\omega)$ is an N-element vector of complex array observation samples, $s_k(\omega)$ is a complex signal amplitude, and $\mathbf{n}(\omega)$ is a vector of complex noise samples. The vector $\mathbf{a}_k(\omega,\phi,\theta)$ is the transfer function between the array and source k located on bearing ϕ at range r and depth z given in $\theta=[r,z]$. We assume that of the Q sources, one is considered to be the TOI while the remaining are considered interference. No assumption is made regarding the relative power levels of the sources when defining which are interference. The CSDM of the array data is formed from M observations as $\mathbf{Y}(\omega)=\Sigma_{l=1}^{M}\mathbf{y}_l(\omega)\mathbf{y}_l^H(\omega)$, where superscript H denotes conjugate transpose. An eigendecomposition of $\mathbf{Y}(\omega)$ is computed as

$$\mathbf{Y}(\omega)=\mathbf{U}(\omega)\mathbf{\Lambda}(\omega)\mathbf{U}^{\mathbf{H}}(\omega), \tag{2}$$

where $\mathbf{\Lambda}(\omega)$ is a diagonal matrix of eigenvalues sorted in descending order, $\sigma_1>\sigma_2\cdots>\sigma_N$, and the columns of the matrix $\mathbf{U}(\omega)=[\mathbf{u}_1(\omega)\mathbf{u}_2(\omega)\cdots\mathbf{u}_N(\omega)]$ are the corresponding eigenvectors of $\mathbf{Y}(\omega)$ which span the column space of the CSDM. The eigenvectors corresponding to the L most significant eigenvalues span the TOI plus interference subspace which includes the Q sources, while the remaining $N-L$ eigenvectors span its' orthogonal complement. Determining which of the L TOI plus interference subspace eigenvectors corresponds predominantly to interfering sources is key to effectively suppressing them.

The ECA method uses beamforming of the eigenvectors of the CSDM as a means of identifying which of the TOI plus interference subspace eigenvectors contain significant contributions from the TOI. A plane-wave beamformer incoherently averaged using K frequency components over the frequency band of the TOI is applied to the eigenvectors as

$$B_i(\phi)=\sum_{l=1}^{K}\left|\mathbf{w}^H(\phi,\omega_l)\mathbf{u}_i(\omega_l)\right|^2, \tag{3}$$

where $\mathbf{w}(\phi,\omega_l)$ is the beamformer weight vector in the direction ϕ at frequency ω_l. For each eigenvector across frequency, $B_i(\phi)$ is computed for ϕ from 0 to 180 deg with beam spacing selected to fully cover bearing space. Therefore (3) computes the average contributions of the sources occupying each eigenvector across frequency.

If we assume that an estimate of the TOI's bearing $\hat{\phi}_o$ is known, for each $B_i(\phi)$ we can compare, over the eigenvector index i, the power at $\hat{\phi}_o$ with the power in adjacent bearings, $\mathbf{\Phi}_A=[\phi_{-J}\cdots\phi_{-1}\phi_1\cdots\phi_J]$, defined as a set of J bearings on either side of $\hat{\phi}_o$. We then compute the power ratio

$$PR_i=\frac{B_i(\hat{\phi}_o)}{\max\limits_{\mathbf{\Phi}_A} B_i(\mathbf{\Phi}_A)}, \tag{4}$$

for each eigenvector index and compare it to a threshold. Those eigenvectors across frequency associated with a PR_i which is below the threshold value are considered to be dominated by interference and are subtracted from the CSDM using the projection

$$\tilde{\mathbf{Y}}(\omega)=\mathbf{P}_\perp(\omega)\mathbf{Y}(\omega)\mathbf{P}_\perp(\omega), \tag{5}$$

at each of the K frequencies. The projection matrix $\mathbf{P}_\perp(\omega)$ is formed from the set of eigenvectors whose PR_i are less than the threshold value as

$$\mathbf{P}_\perp(\omega)=\mathbf{I}-\check{\mathbf{U}}(\omega)\check{\mathbf{U}}^{\mathbf{H}}(\omega), \tag{6}$$

where the columns of $\check{\mathbf{U}}(\omega)$ are the eigenvectors $\mathbf{u}_i(\omega)$ whose PR_i are below the threshold and $\mathbf{I}$ is an identity matrix. The matrix resulting from the product $\check{\mathbf{U}}(\omega)\check{\mathbf{U}}^{\mathbf{H}}(\omega)$ is a projection matrix onto the estimated interference subspace. Therefore, the operation in (5) subtracts from $\mathbf{Y}(\omega)$ that portion of the data which lies in the estimated interference subspace producing the reduced-rank CSDM $\tilde{\mathbf{Y}}(\omega)$. It is important to note that if in the subsequent processing of $\tilde{\mathbf{Y}}(\omega)$ an assumed signal model is employed for the TOI, it may be necessary to also apply $\mathbf{P}_\perp(\omega)$ to the signal model vectors. If the TOI has a component which lies in the estimated interference subspace, it will be removed in (5). Therefore, that component must also be removed from the signal model using $\mathbf{P}_\perp(\omega)$ in order to match the TOI. Only in the unlikely event that the TOI is orthogonal to the interfering sources will this not be applicable.

In this manner, ECA adaptively estimates the interference subspace at each CSDM update and removes it from the CSDM. It is not constrained to define the interference subspace as low rank of a specific dimension. Moreover, ECA allows the estimated interference subspace to be comprised of nonsuccessively ordered eigenvectors. This is important because motion and changing power levels will cause targets

to occupy different eigenvectors as a function of time. Target motion will cause a single target to occupy more than one eigenvector. Additionally, multiple targets with similar power levels may occupy a single eigenvector. Using beam power in (4) as a metric, ECA determines the relative contributions of the TOI versus interference when the TOI and interferers occupy a single eigenvector, thereby providing an informed means of eigenvector removal. Thus, ECA can rapidly adapt to the variability in the TOI plus interference subspace in highly dynamic scenarios.

When the TOI and an interferer cross in bearing, the metric in (4) is no longer useful during the period of time in which they are on the same bearing. We will assume here that some external process is capable of providing notification of the possible onset and completion of a bearing crossing event. Given that, the approach ECA takes in this situation uses the eigenvector removal history for a given time interval just prior to the bearing crossing event. When notification of the onset of a bearing crossing event is given, ECA first determines the percentage of the time each of the eigenvectors across frequency had a PR_i below the threshold for the prior time interval, i.e., the percentage of the time each was subtracted from the CSDM. Those eigenvectors attaining a percentage greater than some set value, for example 50%, are assumed to generate the interference subspace during the crossing event. That is, this fixed set of eigenvectors is subtracted from the CSDM until completion of the bearing crossing event. The underlying assumption, given that the event is of relatively short duration, is that the relative power levels of the sources will not change significantly during the event. Therefore, selecting a fixed set of eigenvectors based on the statistics of their subtraction history will provide adequate interference suppression during the event. After completion of the bearing crossing event, ECA resumes normal operation using (4) to determine the appropriate eigenvectors to subtract from the CSDM.

III. DEMONSTRATION OF ECA METHOD USING SIMULATED DATA

In this section, we present two examples using simulated data which demonstrate the effectiveness of the ECA method in adaptively suppressing interference. We will apply ECA to the problem of source localization using the technique of matched-field processing[4,5] (MFP) in the presence of interference. MFP uses complex multipath propagation models to calculate the transfer function between an acoustic source and a receiving array. Source localization estimates are computed by finding the peak of the Bartlett processor[5] given by

$$\beta(\theta)=\sum_{l=1}^{K}\frac{\mathbf{a}^{\mathbf{H}}(\omega_l,\phi_o,\theta)\mathbf{Y}(\omega_l)\mathbf{a}(\omega_l,\phi_o,\theta)}{\mathbf{a}^{\mathbf{H}}(\omega_l,\phi_o,\theta)\mathbf{a}(\omega_l,\phi_o,\theta)}, \tag{7}$$

where $\mathbf{a}(\omega_l,\phi_o,\theta)$ is the acoustic transfer function for a source located at the range and depth in θ for frequency ω_l, and ϕ_o is the given target bearing provided by a beamformer. The value of θ that produces the maximum value in the above equation is taken as the source range and depth estimate.

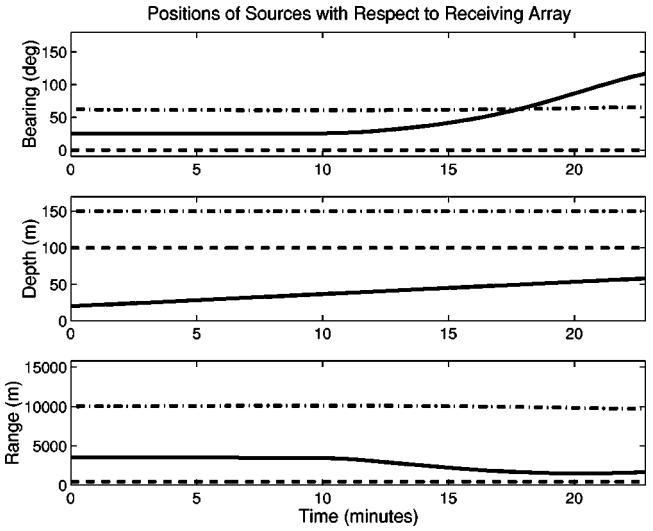


FIG. 1. Positions of sources for first simulation example. TOI is represented by the solid lines.

For the simulations, we assumed a 250-m horizontal array with 5-m spacing between sensors located in the water column of a 300-m-deep range-invariant environment. A normal-mode model[6] was used to compute the acoustic transfer functions of the sources. Both simulations contain approximately 23 min of data consisting of three acoustic sources in the frequency band of 30–150 Hz; one source was considered to be the TOI with the others considered to be interference. The received signal was processed using $K=13$ frequency components evenly spaced across the band. Each CSDM was computed using $M=10$ array observations. For the MFP localization processing, a range search grid spacing of 20 m from 1000 to 6000 m and a depth search grid spacing of 5 m from 10 to 100 m was used. ECA was applied to the output of the array prior to the MFP stage. As discussed in the previous section, in order to properly model the signal transfer function when ECA is employed it was necessary to modify $\mathbf{a}(\omega_l,\phi_o,\theta)$ by premultiplying by the projection matrix $\mathbf{P}_\perp(\omega)$. Modifying (7) to reflect the preprocessing of the data with ECA, and using the projection matrix idempotent property of $\mathbf{P}_\perp(\omega)\mathbf{P}_\perp(\omega)=\mathbf{P}_\perp(\omega)$ gives

$$\tilde{\beta}(\theta)=\sum_{l=1}^{K}\frac{\mathbf{a}^{\mathbf{H}}(\omega_l,\phi_o,\theta)\tilde{\mathbf{Y}}(\omega_l)\mathbf{a}(\omega_l,\phi_o,\theta)}{\mathbf{a}^{\mathbf{H}}(\omega_l,\phi_o,\theta)\mathbf{P}_\perp(\omega_l)\mathbf{a}(\omega_l,\phi_o,\theta)}. \tag{8}$$

The positions of the three sources for the first simulation example are given in Fig. 1. The solid line represents the location of the TOI, while the dashed and dash-dot lines represent the locations of the interferers. Figure 2 shows the average power levels of the sources at the receiving array using the same line types to represent each source. For both simulation examples the average noise power level was 60 dB/Hz, resulting in interference-dominated scenarios. From these figures we see that for approximately the first 10 min the TOI remains on a constant bearing adjacent to an interferer at 0 deg. During this time interval the TOI is approximately 5 dB weaker than the adjacent interferer. After this time, the TOI power begins to increase as it closes in on the array while moving toward broadside. At approximately 18

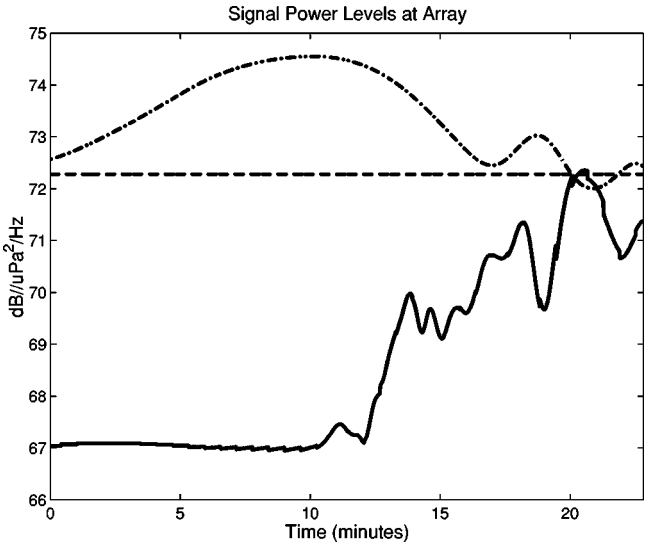


FIG. 2. Power levels of sources for first simulation example. TOI is represented by the solid line.

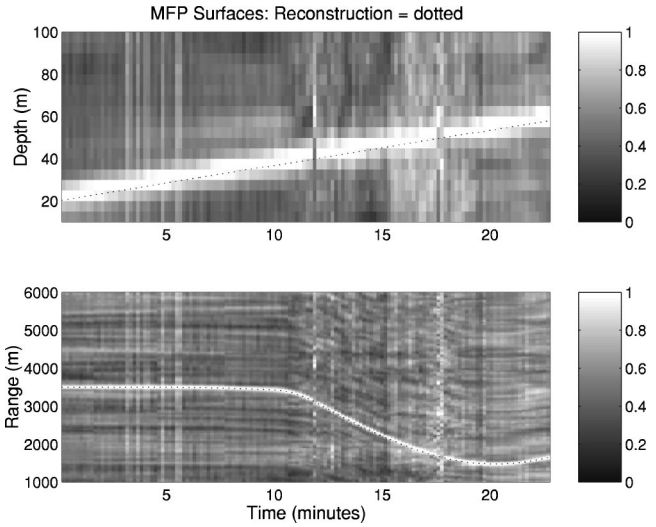


FIG. 4. Range and depth localization surfaces for first simulation example computed with ECA interference suppression applied to the data.

min, the TOI crosses the bearing of the second interferer and continues towards aft endfire of the array.

Figure 3 shows the MFP localization results obtained without using ECA interference suppression, i.e., processing the data using Eq. (7). The plots show slices through the surfaces produced by (7) at the values of range and depth where $\beta(\theta)$ is maximum. That is, at each time update the depth surface shows the slice through depth at the range where $\beta(\theta)$ is maximum and the range surface shows the slice through range at the depth where $\beta(\theta)$ is maximum. Each time slice is normalized for a peak value of 1. We see that over the first 10 min when the TOI is near in bearing to an interferer, no useful localization estimates are produced except for the period between approximately 2 to 4 min, where some intermittently correct results are observed. As the TOI begins to move away in bearing from this interferer, correct localization estimates are obtained until the TOI approaches the bearing of the second interferer. The localization estimates again become degraded in the vicinity where the TOI crosses the bearing of the second interferer. Once the TOI moves away from the bearing of the second interferer, correct localization estimates are once again produced.

Processing this same data set by applying ECA interference suppression prior to computing MFP localization estimates [Eq. (8)] produced the results shown in Fig. 4. ECA was applied using a power ratio threshold of 0.9 and $J=15$ adjacent bearings when computing Eq. (4). Using ECA, we see that accurate localization estimates are obtained for the entire data run. ECA effectively suppresses the first interferer during the first 10 min and the second interferer over the interval where the TOI crosses its bearing. Overall, the surfaces in Fig. 4 more clearly display the trajectory of the TOI, due to the effective suppression of the interferers by ECA, than do those of Fig. 3.

In Fig. 5, we show the beamformer output [Eq. (3)] for the four principal eigenvectors of the CSDM across frequency, plotted in order of significance, for this data set. A set of 101 cosine-spaced beams from 0 to 180 deg was used.

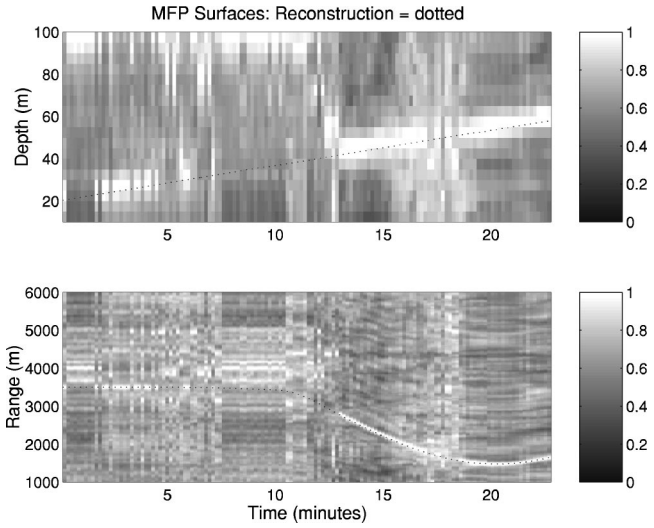


FIG. 3. Range and depth localization surfaces for first simulation example computed without applying ECA interference suppression to the data.

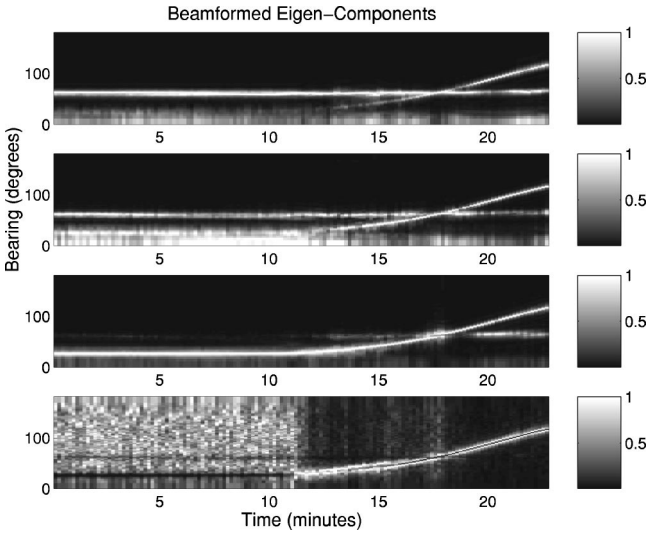


FIG. 5. Beamformer output for the four principal eigenvectors of first simulation example. Outputs plotted in order of eigenvector significance.

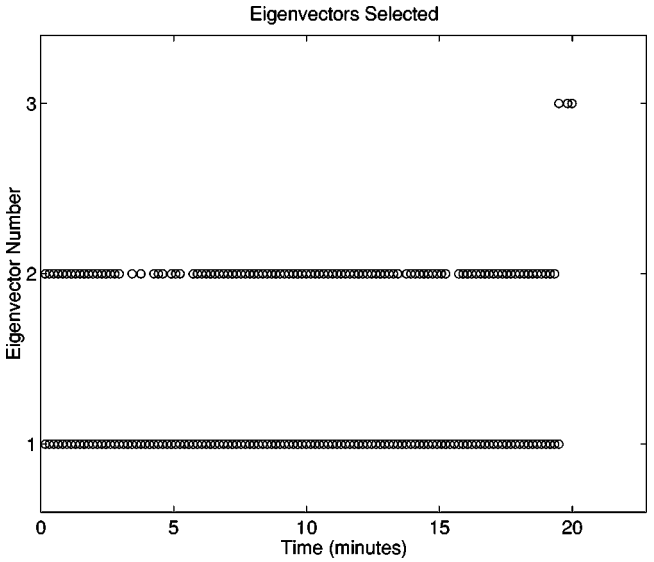


FIG. 6. Eigenvector numbers identified by ECA for subtraction from CSDM for first simulation example.

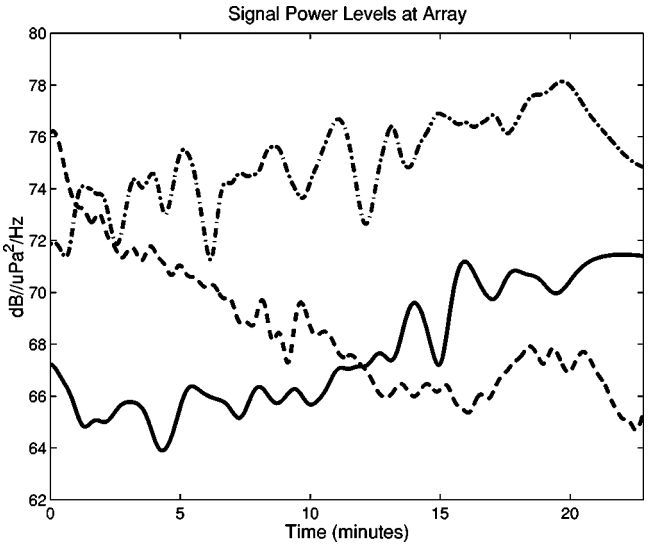


FIG. 8. Power levels of sources for second simulation example. TOI is represented by the solid line.

The outputs at each time update are normalized to a peak value of 1. The relative amplitude of the output observed beyond the third eigenvector was insignificant; therefore, only the first three eigenvectors were considered by ECA. With three sources it would be expected that the TOI plus interference subspace would occupy the first three eigenvectors. The fourth eigenvector contains a very weak TOI contribution spread to it due to target motion. From the plots we can observe the relative contributions of each of the sources that are contained in each of the eigenvectors. The fluctuations of the contributions as a function of time are also evident. We observe that the first eigenvector predominantly contains the second interferer for the first 19 min after which it contains a fluctuating mix of the interferer and the TOI. The second eigenvector consists of a varying mix of all three sources, with the TOI becoming dominant in the latter half of the run. The third eigenvector is dominated by the TOI with the second interferer showing a significant contribution around 20 min. Figure 6 depicts at each time update the eigenvectors that ECA determined to be dominated by interference and were subtracted from the CSDM. From the beginning of the run until 19 min, the first eigenvector is always selected, the second eigenvector is selected the majority of the time, and the third eigenvector is never selected. After 19 min until the end of the run, only the third eigenvector is selected for a brief period around 20 min which coincides with the increase in the second interferer's energy.

The positions of the three sources for the second simulation example and their associated average power levels are given in Figs. 7 and 8, with the solid line representing the TOI. In this example, the TOI is sandwiched closely between the two interferers in bearing. The power of the interferer on the lower bearing increases with time with a trend similar to the TOI, while the interferer on the upper bearing has decreasing power with time.

MFP localization results obtained without using ECA

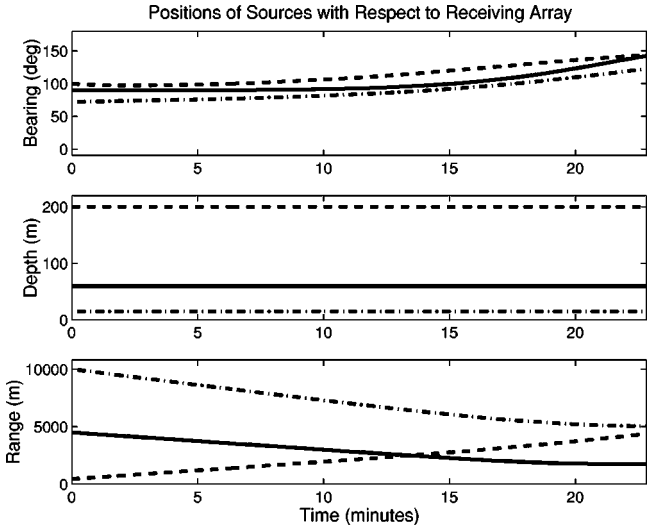


FIG. 7. Positions of sources for second simulation example. TOI is represented by the solid lines.

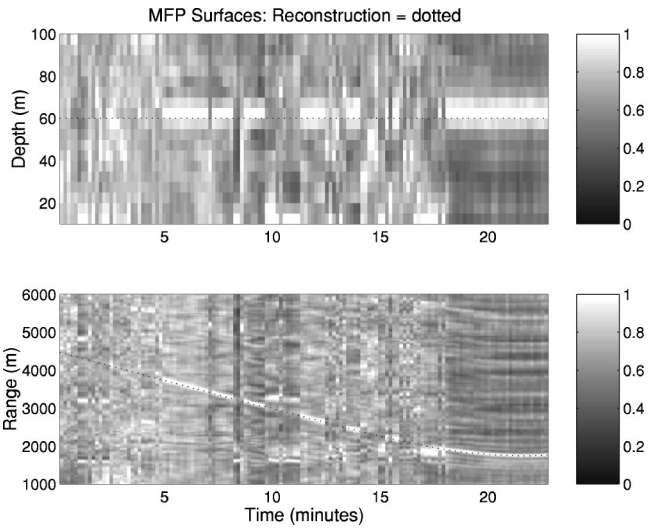


FIG. 9. Range and depth localization surfaces for second simulation example computed without applying ECA interference suppression to the data.

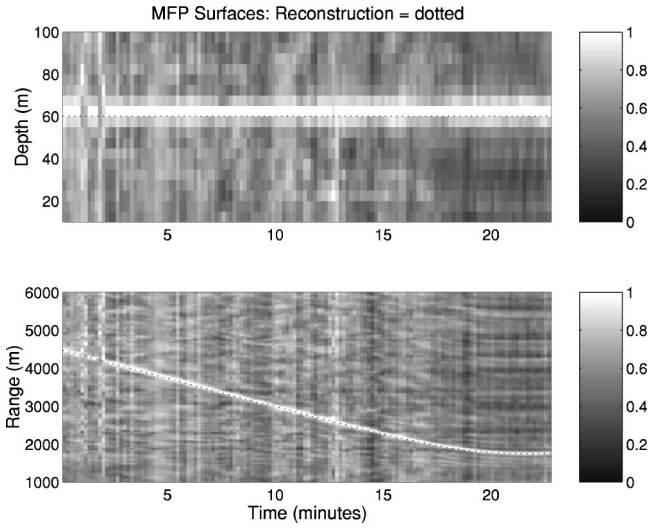


FIG. 10. Range and depth localization surfaces for second simulation example computed with ECA interference suppression applied to the data.

are shown in Fig. 9. During the first 5 min while the TOI is close in bearing to the upper interferer, no useful localization estimates are produced. Between 5 to 10 min as the upper interferer's power weakens, some reasonable localization results are produced. The TOI begins to approach the stronger, lower interferer in bearing between 10 to 18 min, resulting in degraded localization estimates during this time period. After 18 min, the TOI moves away from the lower interferer in bearing back towards the weaker, upper interferer and correct localization estimates are produced.

Figure 10 shows the localization results obtained processing the second example using ECA prior to MFP. Once again, accurate localization estimates are obtained for the entire data run when ECA is applied to the data. The beamformer output for the four principal eigenvectors of the CSDM across frequency are shown in Fig. 11 for this data set. Again, no significant signal contributions were present beyond the third eigenvector, so only the first three were considered by ECA. We observe that the principal eigenvector is dominated by the lower interferer for the entire run. The second eigenvector is a mix of the three sources, but predominantly contains the upper interferer for the first half of the run, and is then dominated by the TOI for the second half. The third eigenvector is dominated by the TOI for the first half of the run, and then the upper interferer for the second half. Threrefore, what we are observing here is a hierarchical swap of the TOI and upper-interferer occupied eigenvectors around the midpoint of the data run. Figure 12 depicts the eigenvectors ECA selected for subtraction from the CSDM at each time update. We observe that ECA adapted to the hierarchy swap near the midpoint of the run, after which time the estimated interference subspace was generated by nonsuccessively ordered eigenvectors.

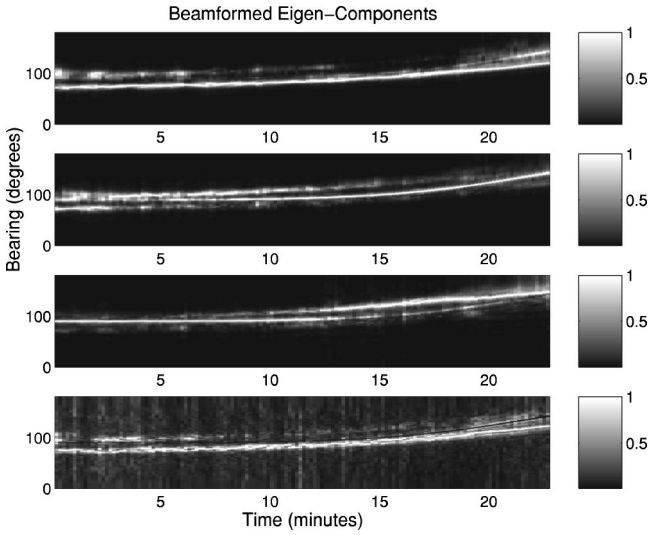


FIG. 11. Beamformer output for the four principal eigenvectors of second simulation example. Outputs plotted in order of eigenvector significance.

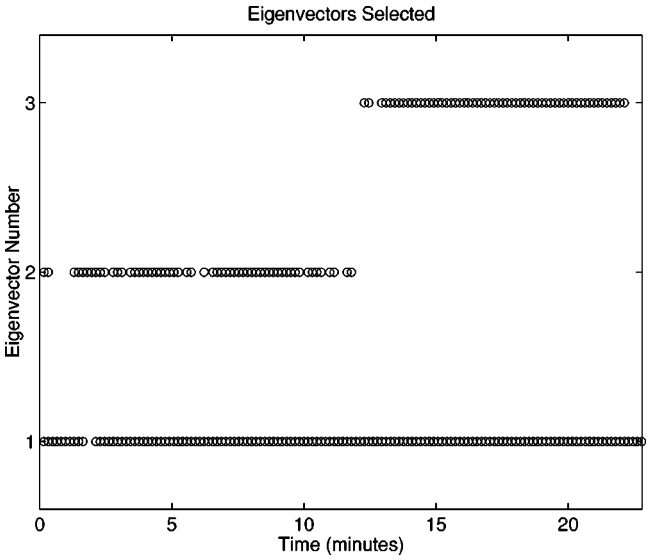


FIG. 12. Eigenvector numbers identified by ECA for subtraction from CSDM for second simulation example.

IV. CONCLUSION

This paper presented the ECA method of adaptive interference suppression. ECA builds on the PCI method by utilizing bearing information to determine which eigenvectors of the CSDM are dominated by interference. Once identified, these eigenvectors are subtracted from the CSDM to suppress the interfering signals. Simulation examples were presented to demonstrate the effectiveness of ECA in rapidly adapting to variations in the interference subspace.

ACKNOWLEDGMENT

This work was supported by the Office of Naval Research, Dr. John Tague, ONR 321US.

[1] I. P. Kirstens and D. W. Tufts, "On the probability density of signal-to-noise ratio in an improved adaptive detector," in Proc. IEEE ICASSP-85, Tampa, FL, 1985, pp. 572–575.
[2] I. Kirsteins and D. Tufts, "Adaptive detection using low rank approximation to a data matrix," IEEE Trans. Aerosp. Electron. Syst. **30**, 55–67 (1994).
[3] S. M. Kogon, "Robust adaptive beamforming for passive sonar using eigenvector/beam association and excision," in Proceedings of 2002 Sensor Array and Multichannel Signal Processing Workshop, Lexington, MA, 2002.

[4] H. P. Bucker, ''Use of calculated sound fields and matched-field detection to locate sound sources in shallow water,'' J. Acoust. Soc. Am. **59**, 368–373 (1976).
[5] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, ''An overview of matched field methods in ocean acoustics,'' IEEE J. Ocean. Eng. **18**, 401–424 (1993).
[6] M. B. Porter, ''The KRAKEN normal mode program,'' Tech. Rep. 6920, Naval Research Laboratory, 1992.

Experimental validation of regularized array element localization

Stan E. Dosso[a)]
School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia V8W 3P6, Canada

Nicole E. B. Collison, Garry J. Heard, and Ronald I. Verrall
Defence Research and Development Canada (Atlantic), Dartmouth, Nova Scotia B2Y 3Z7, Canada

(Received 18 April 2003; revised 8 January 2004; accepted 5 February 2004)

This paper examines and validates regularized inversion for array element localization (AEL) by quantitative comparison of inversion results to direct measurements of receiver positions for a full-scale AEL survey. Regularized AEL treats both receiver and source positions as unknown parameters in a ray-based inversion; prior information on source/receiver positions, inter-receiver spacing in depth, and/or a smooth array shape can be included, subject to statistically fitting the acoustic data. Uncertainties in the recovered receiver positions are estimated via Monte Carlo appraisal. To study this approach, a specially stabilized, two-dimensional receiver array and a series of impulsive sources (imploding glass light bulbs) were deployed from shore-fast (motionless) Arctic sea ice. Sources and recordings were not synchronized in time, so AEL inversions are based on relative arrival times. Receiver positions were measured to an uncertainty of ~5 cm in each dimension [9 cm in three dimensions (3D)] using nonacoustic (optical) methods. Average AEL errors (difference between measured receiver positions and inversion results) of 13 cm in depth, 27 cm in the horizontal, and 30 cm in 3D, as well as good agreement between the measured errors and estimated AEL uncertainties validate the regularized approach and provide benchmarks for acoustic AEL. Receiver-position errors are quantitatively investigated as a function of the number of sources, source-position errors, and different regularizations. [DOI: 10.1121/1.1701897]

PACS numbers: 43.60.Pt, 43.30.Xm [WLS] Pages: 2129–2137

I. INTRODUCTION

Advanced array processing methods, such as beamforming[1] and matched-field processing and inversion,[2–5] require accurate knowledge of the position of individual receivers in the array. For example, a general rule to achieve an array processing gain within 1 dB of the theoretical limit requires receiver positions be known within a tenth of a wavelength at the frequency of interest.[6,7] However, at-sea array deployment is not an exact procedure, and an acoustic survey is often required after deployment to localize the receivers, a procedure known as array element localization (AEL). AEL surveys typically consist of measuring the travel times of acoustic signals at the receivers from a number of sources positioned around the array. Given knowledge of the ocean sound-speed profile, the travel-time data can be inverted for estimates of the receiver positions. Simple surveys based on expendable impulsive sources (e.g., explosive charges or imploding glass light blubs[8,9]) without synchronization in time between source transmissions and the receiver recording system provide a practical approach to AEL that is widely used.[9–12]

Ideally, AEL inversion should address all (significant) sources of error in the acoustic survey, and incorporate available *a priori* information in addition to the measured data. For instance, although source positions are usually treated as known in AEL, in practice, errors in these positions are often significant and represent the limiting factor.[13–16] Regularized AEL[14] addresses this limitation by treating both source and receiver positions as unknown parameters (with prior estimates and uncertainties) in a ray-tracing based inversion. An unknown bias for the sound-speed profile is included to account for calibration errors in water-column measurements, and prior estimates of the inter-receiver spacing in depth can also be incorporated. For arrays that are expected to be essentially straight, a regularization can be formulated for the smoothest array shape subject to statistically fitting the acoustic data. This provides the simplest shape consistent with the data, with no deviations from a smooth array that are not resolved by the data (a common consequence of overfitting noisy data).[14] The uncertainties of the recovered receiver positions are estimated using nonlinear Monte Carlo appraisal.[14] Regularized AEL has been applied to fixed horizontal arrays,[14,17] moored vertical arrays,[17,18] towed arrays,[15,19] and freely drifting sonobuoy fields.[16] Note, however, that independent measurements of receiver positions were not available in any of these applications to validate the AEL results, nor does it appear that any such ''ground-truthing'' of AEL methods has been reported to date.

The goal of this paper is to examine and validate regularized AEL by quantitative comparison of inversion results to direct measurements of receiver positions for a full-scale AEL survey. High-precision nonacoustic measurement of receiver positions required novel approaches. The experiment was carried out from shore-fast Arctic sea ice, which pro-

[a)]Electronic mail: sdosso@uvic.ca

vided a motionless platform. The array consisted of a horizontal line array (HLA) of geophones on the ice surface and a vertical line array (VLA) of hydrophones below the ice. A special six-point suspension system was employed to constrain the VLA in three dimensions so that it did not move with currents during the experiment.[20,21] Receiver positions were measured using a laser survey for the HLA and a remotely operated vehicle (ROV) for the VLA. Light-bulb sources were deployed at a number of positions around the array. The regularized AEL results are compared to the measured receiver locations, and the resulting errors are compared to the uncertainties estimated using Monte Carlo appraisal. Several cases are examined to quantitatively investigate the effect of the number of sources, treatment of source-position errors, and different forms of *a priori* information (i.e., different regularizations). The results presented here represent an important benchmark for the accuracy of acoustic AEL surveys of this type.

II. REGULARIZED AEL INVERSION ALGORITHM

The theoretical background and implementation for regularized AEL inversion has been reported previously.[14,15] This section provides an updated overview of the approach with several new elements; more complete treatments of the general theory of regularized inversion can be found elsewhere.[22–24] For AEL inversion, the relative travel times from sources to receivers can be written in vector form as

$$\mathbf{t}=\mathbf{T}(\mathbf{m})+\mathbf{n}. \tag{1}$$

In Eq. (1), the model $\mathbf{m}$ of unknown parameters consists of the three-dimensional (3D) position (x,y,z) of each receiver, the 3D position and transmission instant of each source, and a bias correction term for the sound-speed profile. The forward mapping $\mathbf{T}(\mathbf{m})$ represents the acoustic travel times along ray paths from source to receiver, and contains all of the physics and geometry of the problem. Finally, $\mathbf{n}$ represents the data errors (noise), with the assumption that the error n_i on datum t_i is due to an independent, Gaussian-distributed random process with zero mean and standard deviation σ_i. The inverse problem of determining $\mathbf{m}$ from $\mathbf{t}$ is functionally nonlinear. However, a local linearization can be obtained by expanding $\mathbf{T}(\mathbf{m})=\mathbf{T}(\mathbf{m}_0+\delta\mathbf{m})$ in a Taylor series to first order about an arbitrary starting model $\mathbf{m}_0$. Rearranging terms, this expansion can be written

$$\mathbf{J}\mathbf{m}=\mathbf{t}-\mathbf{T}(\mathbf{m}_0)+\mathbf{J}\mathbf{m}_0\equiv\mathbf{d}, \tag{2}$$

where $\mathbf{d}$ represents modified data defined in terms of known quantities and $\mathbf{J}$ is the Jacobian matrix of partial derivatives (sensitivities) $J_{ij}=\partial T_i(\mathbf{m}_0)/\partial m_j$ (an efficient formulation for computing ray travel times and analytic derivatives, including direct, turning, and reflected paths, is given in Refs. 15 and 16). Equation (2) represents a linear inverse problem which can be solved for $\mathbf{m}$ as described in the following. Since nonlinear terms are neglected, the inversion must be repeated iteratively while monitoring for convergence (discussed in the following).

Treating both source and receiver positions as unknown leads to an ill-posed problem that is not amenable to a standard least-squares solution.[22] The method of regularized inversion[14,22–24] provides a powerful approach to ill-posed problems based on incorporating physical *a priori* information to define a well-posed problem. Other approaches, such as singular-value decomposition and Levenberg–Marquardt inversion, are also applicable and provide a mathematically stable solution, but without the physical relevance imposed by prior information (discussed for AEL in Ref. 14). For regularized AEL inversion, three different forms of prior information can be imposed by including three regularization terms together with the data misfit in an objective function ϕ to be minimized:

$$\phi=|\mathbf{G}(\mathbf{J}\mathbf{m}-\mathbf{d})|^2+\mu_1|\mathbf{H}_1(\mathbf{m}-\hat{\mathbf{m}}_1)|^2+\mu_2|\mathbf{H}_2(\mathbf{m}-\hat{\mathbf{m}}_2)|^2+\mu_3|\mathbf{H}_3(\mathbf{m}-\hat{\mathbf{m}}_3)|^2. \tag{3}$$

In Eq. (3), the first term represents the χ^2 data misfit for the linear problem (given Gaussian data errors), with $\mathbf{G}=\mathrm{diag}[1/\sigma_i]$, and the remaining terms represent regularizations. The variables μ_1, μ_2, and μ_3 represent trade-off parameters (Lagrange multipliers) which determine the relative importance of the misfit and prior information in the minimization.

The first regularization term in Eq. (3) can be used to apply prior estimates for the model parameters, which in AEL applications consist of approximate knowledge of source and/or receiver positions from the deployment procedure. In this case, $\hat{\mathbf{m}}_1$ contains the prior estimates and the regularization matrix $\mathbf{H}_1$ weights the estimates according to their uncertainties,

$$\mathbf{H}_1=\mathrm{diag}[1/u_j], \tag{4}$$

where u_j represents the standard deviation of an assumed Gaussian uncertainty distribution for jth parameter.

The second regularization term in Eq. (3) can be used to apply prior information about the one-dimensional (1D) separation between receivers along a coordinate axis, such as depth or horizontal offset [specifying two-dimensional (2D) or 3D separations involves nonlinear constraints which cannot be included in linearized inversion]. To accomplish this, $\hat{\mathbf{m}}_2$ contains the expected 1D offsets v_j of the array along a coordinate axis to receiver j. Regularization matrix $\mathbf{H}_2$ consists of a (nonsquare) piecewise bidiagonal matrix (i.e., there are breaks between rows corresponding to the x, y, and z coordinates) with nonzero entries on the jth row given by

$$\mathbf{H}_2=\mathrm{diag}\left[\frac{-1}{v_{j+1}-v_j},\frac{1}{v_{j+1}-v_j}\right]. \tag{5}$$

This matrix weights prior information for closely spaced receivers more strongly than for widely spaced receivers, and has not been applied previously in regularized AEL inversion.

The third regularization term can apply the *a priori* expectation that the array shape is well approximated by a smooth function (i.e., the array shape involves minimal curvature or changes in direction). This can be applied using $\hat{\mathbf{m}}_3=0$ and $\mathbf{H}_3$ consisting of a (piece-wise) tridiagonal matrix with nonzero entries on the jth row given by

$$\mathbf{H}_3 = \mathrm{diag}\left[\frac{-1}{(w_{j+1}-w_j)^2}, \frac{w_{j+2}-w_j}{(w_{j+2}-w_{j+1})(w_{j+1}-w_j)^2}, \frac{-1}{(w_{j+2}-w_{j+1})(w_{j+1}-w_j)}\right], \tag{6}$$

where w_j represents the expected 3D offset along the array to receiver j. Each row of $\mathbf{H}_3$ in Eq. (6) represents a discrete approximation to the second derivative operator $\partial^2/\partial w^2$ (note that for arrays with equally spaced receivers this simplifies to $\mathbf{H}_3 \propto \mathrm{diag}[-1,2,-1]$ which has been used previously[14–16]). Hence, $|\mathbf{H}_3\mathbf{m}|^2$ provides a measure of the total curvature of the array, and the regularization produces the minimum-curvature or smoothest (simplest) shape consistent with the acoustic data. The minimum-curvature regularization can be applied in one, two, or three dimensions, as appropriate.

Minimizing objective function (3) with respect to $\mathbf{m}$ yields the solution

$$\begin{aligned}\mathbf{m} = [&\mathbf{J}^T\mathbf{G}^T\mathbf{G}\mathbf{J} + \mu_1\mathbf{H}_1^T\mathbf{H}_1 + \mu_2\mathbf{H}_2^T\mathbf{H}_2 \\ &+ \mu_3\mathbf{H}_3^T\mathbf{H}_3]^{-1}\cdot[\mathbf{J}^T\mathbf{G}^T\mathbf{G}\mathbf{d} + \mu_1\mathbf{H}_1^T\mathbf{H}_1\hat{\mathbf{m}}_1 \\ &+ \mu_2\mathbf{H}_2^T\mathbf{H}_2\hat{\mathbf{m}}_2 + \mu_3\mathbf{H}_3^T\mathbf{H}_3\hat{\mathbf{m}}_3].\end{aligned} \tag{7}$$

Note that the second and third regularization terms are optional, and can be omitted if the appropriate prior information is not available by setting μ_2 and/or μ_3 to zero.

Implementation of the AEL algorithm consists of an iterative application of the regularized inversion (7), initiated from a starting model coinciding with the prior parameter estimates. Convergence of the algorithm is based on two criteria. First, the measured data must be fit such that the χ^2 misfit achieves its expected value of $\langle\chi^2\rangle = N$ for N data.[24] Note that although Eq. (7) is derived based on the linearized inverse problem, convergence must be judged in terms of the nonlinear misfit

$$\chi^2 = |\mathbf{G}(\mathbf{T}(\mathbf{m}) - \mathbf{t})|^2. \tag{8}$$

Second, a stable solution must be obtained such that the root-mean-square change in receiver positions between iterations is small compared to the expected accuracy of the solution (<1 cm for the present application). A practical aspect of implementing the inversion involves assigning values to the trade-off parameters, μ_1, μ_2, and μ_3. An effective procedure for two regularization terms is described in Refs. 14–16; extension to three regularizations is straightforward.

Uncertainty estimates for the regularized AEL solution are obtained using nonlinear Monte Carlo appraisal.[14–17,24] The source and receiver positions determined via inversion of measured data are assumed to define the true positions for a synthetic AEL problem, and acoustic travel-time data are computed for these positions. Subsequently, a large number of independent inversions of the synthetic data are carried out, each with different random errors applied to the data and to the prior source/receiver position estimates. The errors are drawn from Gaussian distributions with standard deviations corresponding to the estimated uncertainties for the measured data and priors. Standard deviations about the true positions are computed from the ensemble of synthetic inversion results to provide parameter uncertainty estimates.

The above-mentioned Monte Carlo procedure can be used to estimate localization errors in both an absolute sense (relative to the fixed, geographic coordinate system) and in a relative sense (relative to array-based coordinates). For relative error estimates, the effects of rigid-body translations and rotations of the individual inversion estimates (relative to the true positions) must be removed prior to computing error statistics. Translations are removed by aligning the 3D centroid of the estimated sensor positions with that of the true positions. To minimize the effects of rotations between a set of N estimated sensor positions (x_i, y_i) and reference positions (X_i, Y_i), an optimal rotation can be computed as follows. Measuring angles with respect to the x axis, rotated estimated positions are defined

$$\tilde{x}_i = r_i\cos(\theta_i + \psi), \quad \tilde{y}_i = r_i\sin(\theta_i + \psi), \tag{9}$$

where $r_i = [x_i^2 + y_i^2]^{1/2}$, $\theta_i = \tan^{-1}(y_i/x_i)$, and ψ is the rotation angle to be determined. The l_2 error norm between the rotated and reference positions is defined

$$E = \sum_{i=1}^{N} \{[r_i\cos(\theta_i + \psi) - X_i]^2 + [r_i\sin(\theta_i + \psi) - Y_i]^2\}. \tag{10}$$

Setting $\partial E/\partial\psi = 0$ to minimize the error leads (after some algebra) to

$$\psi = \tan^{-1}\left[\frac{\Sigma_{i=1}^N r_i(Y_i\cos\theta_i - X_i\sin\theta_i)}{\Sigma_{i=1}^N r_i(X_i\cos\theta_i - Y_i\sin\theta_i)}\right]. \tag{11}$$

Applying the rotation angle given in Eq. (11) to the estimated sensor positions, Eq. (9), produces the position estimates with minimal error due to rotation.

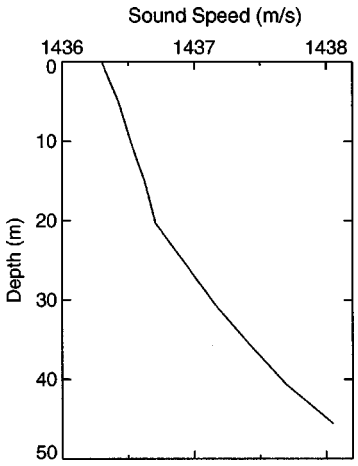


FIG. 1. Sound speed profile measured at experiment site.

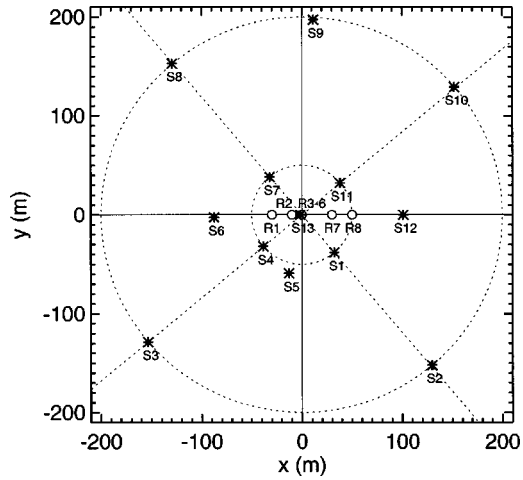


FIG. 2. Plan view of AEL survey. Open circles indicate receiver locations, with the VLA at the origin; asterisks indicate source positions.

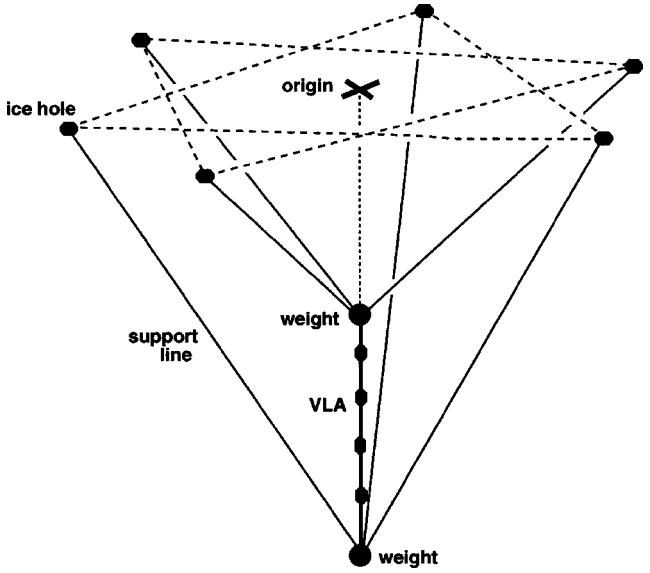


FIG. 3. Stable six-point suspension system used to prevent VLA movement. The triangular arrangement supporting each end of the VLA involves three ice holes at distances of 40, 40, and 65 m from the origin.

III. EXPERIMENT AND DATA

The goal of this work is to compare the results of regularized AEL to actual receiver positions for a full-scale AEL survey. This requires high-precision (nonacoustic) measurements of receiver positions and a means of constraining the receivers so they did not move during the survey. To accomplish this, the experiment was performed on shore-fast ice of the Lincoln Sea north of Ellesmere Island, Canada, which provided a stable, motionless platform. The experiment site was on a large expanse of annual ice which ranged in thickness from 1.4 to 1.7 m, with a water depth of ~45 m. The weather during the experiment was clear and cold (−30 to −40 °C) with no wind, conditions that typically result in minimal ambient noise due to ice-cracking[25] or pressure ridging events.[26] The ocean sound-speed profile was measured using a CTD (conductivity, temperature, depth) cast, and ice properties were determined using a hammer seismic survey.[27] The water-column sound speed, shown in Fig. 1, is upward-refracting over the entire 45 m depth, providing a clearly defined and unique direct ray path from source to receiver.

The receiver array consisted of an HLA of four vertical-component geophones on the ice and a VLA of four hydrophones suspended below the ice. The array is illustrated in plan view in Fig. 2 and receiver positions are given in Table I. The origin of the coordinate system for the experiment was taken to be directly above the intended position of the VLA, with the x axis aligned along the HLA and z defined to be positive downward from sea level. To ensure that the VLA receivers did not move with currents, a special suspension system was employed to constrain the VLA in three dimensions. This system supported each end of the VLA by three lines under tension in a triangular arrangement, as shown in Fig. 3. Weights of 16 and 28 kg were attached to the top and bottom of the VLA. This arrangement is comparable to supporting each end of the VLA by an inverted tripod, and a large lateral force is required to deflect the array by even a small amount (i.e., VLA deflections due to currents are assumed negligible). The potential difficulty is that there are in fact four lines supporting the bottom weight (three support lines and the VLA), and one of these could be slack. To counter this, an extensible member was included in series with the VLA. The spacing between the two ends of the VLA was adjusted so that the tension in the array was high enough to keep it straight, but not so high that the lower weight was lifted. An analysis of the stability of the suspension system is given in Refs. 20 and 21.

The positions of the HLA receivers were measured using a laser surveying instrument (geodimeter), which provided

TABLE I. Summary of receiver positions and their uncertainties from geodimeter and ROV measurements, and of measured AEL errors. Sensor type g/ph refers to a geophone and h/ph to a hydrophone.

Receiver	Sensor type	Measured positions (m)	Measurement uncertainties (cm)	AEL errors (cm)	
				Absolute	Relative
$R1$	g/ph	(−30.00,0.00,−0.15)	(5,5,5)	(10,61,6)	(30,5,9)
$R2$	g/ph	(−10.00,0.00,−0.15)	(5,5,5)	(43,86,15)	(24,11,10)
$R3$	h/ph	(0.33,0.13,31.75)	(7,7,5)	(36,53,31)	(17,18,26)
$R4$	h/ph	(0.32,0.12,32.75)	(7,7,5)	(35,56,12)	(15,12,12)
$R5$	h/ph	(0.31,0.12,33.26)	(7,7,5)	(35,54,11)	(16,10,16)
$R6$	h/ph	(0.18,0.03,38.67)	(7,7,5)	(42,47,40)	(23,13,17)
$R7$	g/ph	(30.00,0.00,−0.15)	(5,5,5)	(21,94,5)	(40,25,12)
$R8$	g/ph	(50.00,0.00,−0.15)	(5,5,5)	(9,99,7)	(25,12,9)

TABLE II. Summary of light-bulb source positions and uncertainties.

Source	Position (m)	Position uncertainties (m)
*S*1	(32.32,−38.14,40.00)	(1,1,0.5)
*S*2	(129.64,−152.22,40.00)	(1,1,0.5)
*S*3	(−153.36,−128.60,35.00)	(1,1,0.5)
*S*4	(−38.63,−32.10,40.00)	(1,1,0.5)
*S*5	(−12.95,−59.20,40.00)	(1,1,0.5)
*S*6	(−87.71,−2.68,40.00)	(1,1,0.5)
*S*7	(−32.27,38.22,40.00)	(1,1,0.5)
*S*8	(−129.42,153.06,40.00)	(1,1,0.5)
*S*9	(11.45,197.37,40.00)	(1,1,0.5)
*S*10	(152.01,129.49,40.00)	(1,1,0.5)
*S*11	(37.99,32.23,40.00)	(1,1,0.5)
*S*12	(101.00,0.00,40.00)	(1,1,0.5)
*S*13	(−1.50,0.00,10.00)	(0.5,0.5,0.25)

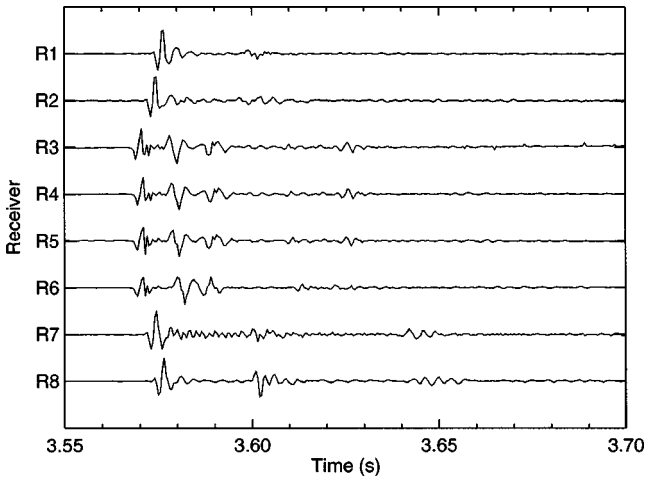


FIG. 4. Example of the recorded time series at the receiver array due to light-bulb source *S*9 in the AEL survey (time scale origin is arbitrary).

positioning accuracy to an estimated uncertainty of 5 cm. The positions of the VLA receivers were measured relative to a 5 kg lead ball lowered on a light kevlar line through the ice at the coordinate origin. The positions of the VLA hydrophones relative to the ball were estimated using the video display relayed from a ROV. The video indicated when the lead ball was aligned at the same depth as each hydrophone, providing depth measurements to an estimated uncertainty of 5 cm (this heavy ball on a light line will have negligible deflection due to currents). The horizontal positions of the hydrophones were estimated from the offset to the lead ball, using the ball's 10 cm diameter as a reference. This provided horizontal positions to within an estimated radial error of 10 cm. Assuming the radial error comprises equal, independent errors in x and y, an uncertainty of 7 cm was assigned to these coordinates. The receiver-position uncertainties are summarized in Table I.

The acoustic sources used for the AEL survey were glass light bulbs imploded under hydrostatic pressure in the water column.[8,9] To control the implosion depth, the light bulbs were attached to a 2 kg weight and lowered on a line through the ice to the desired depth; a messenger weight dropped along the line burst the bulb on impact. Three light-bulb sources were deployed at each of 13 positions around the receiver array, as shown in Fig. 2 and given in Table II. The source depths were 40 m at all but two sites. At source position *S*3 the water depth was less than 40 m, so the light bulbs were imploded at 35 m depth. Source position *S*13 was directly above the VLA, and, to avoid tangling lines, the light bulb was imploded at only 10 m depth. The positions of the source deployment holes were surveyed to high precision using the geodimeter; however, the actual position of the light bulbs at depth can vary due to currents in the water. The uncertainties in the measured source positions were estimated to be 1 m in x and y and 0.5 m in z for all sources except *S*13, where these uncertainties were halved due to the shallower deployment depth. Note that although the source position errors are treated as independent, there may be a systematic component due to currents deflecting all sources in the same direction. Systematic errors in source positions lead to a systematic errors in AEL receiver positions. However, such errors do not effect the relative receiver positions, which is the primary measure of AEL performance. The estimated source-position uncertainties are included in Table II. Note that it would be difficult to achieve the high-precision position measurements given in Tables I and II working in open waters.

Figure 4 shows an example of the recorded time series due to light-bulb source *S*9, digitized at a sampling rate of 2 kHz. The first strong impulsive arrival is clearly identified as the direct acoustic wave, with a well-defined first break and wave form indicative of a single arrival (i.e., uncomplicated by multipath, as expected given the simple form of the sound-speed profile in Fig. 1). The direct arrival is followed by a series of water-column multiples and general reverberation that decrease in amplitude with time. Note that the surface-reflected wave (phase-inverted arrival approximately 0.1 s after the direct arrival) is present at the VLA receivers $(R3-R6)$, but not at the surface-mounted HLA receivers.

The data set used here for AEL consists of the direct arrival times averaged over the three recordings for each source position. Calculating this average required that the three sets of arrival times for each source position be optimally aligned in relative time, given that source instants are not known to provide an absolute time frame. This was accomplished using a grid search to determine the two time shifts that minimized an objective function consisting of the summation of the standard deviation about the mean over all arrivals.[14] Figure 5(a) shows the mean travel-time data for all sources and receivers. Since only relative travel times were measured, the data for each source are shown relative to the absolute travel time expected for receiver $R1$, as computed using ray-tracing and the measured positions for this receiver and all sources. Figure 5(b) shows the standard deviations about the mean. Most standard deviations are $\lesssim$0.2 ms; however, a small number are notably larger. In particular, the relatively large standard deviations obtained for source position *S*13 (directly above the VLA) are due to the shallow (10 m) source depth used here. Light-bulb implosions at shallow depths produce weaker acoustic signals with less high-frequency energy, resulting in poorly defined arrival times. These large errors are unfortunate, as source position *S*13

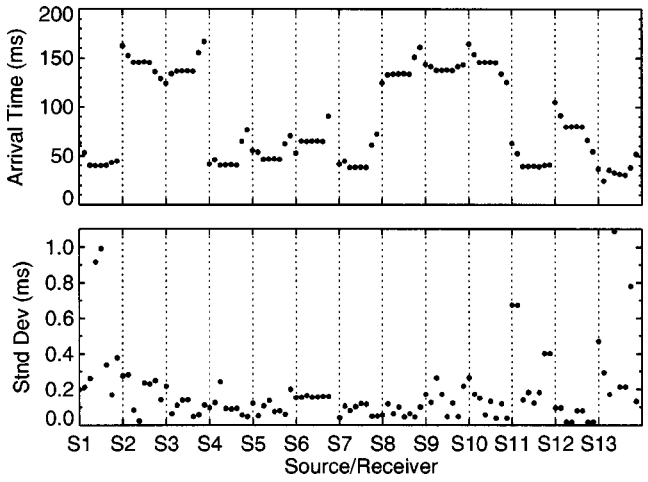


FIG. 5. Data collected in AEL survey: (a) mean arrival times and (b) standard deviations computed from three source deployments at each source location. Sources $S1-S13$ are identified on abscissa, and receivers are ordered $R1-R8$ for each source. The acoustic arrival could not be clearly identified for source $S6$, receiver $R8$ and is not included in the plot.

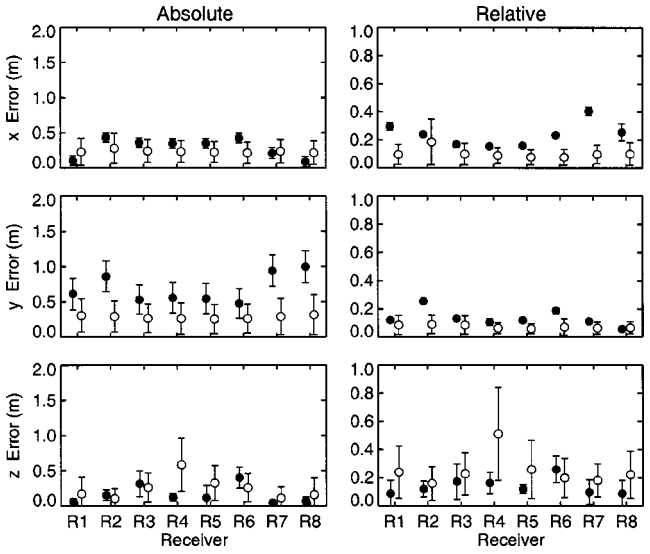


FIG. 6. Receiver localization errors from comparison with measured positions (closed circles) and Monte Carlo analysis (open circles) with prior source location uncertainties of 1 m in x and y and 0.5 m in z. The left column shows absolute errors; the right column shows relative errors.

provides important information constraining the VLA receiver depths. In practice, the small-scale fluctuations in the calculated standard deviations in Fig. 5(b) are not meaningful given the small sample size (three measurements). Hence, the standard deviations used in inversion were set to a fixed value of $\sigma=0.2$ ms for all data with standard deviations <0.3 ms. For data with standard deviations >0.3 ms, the actual standard-deviation values were used.

IV. RESULTS

This section applies the regularized algorithm developed in Sec. II to the AEL survey described in Sec. III, and compares the inversion results for receiver positions to the direct (nonacoustic) measurements. Several cases are considered to quantitatively study various aspects of the regularized inversion.

In the first case, travel-time data from all 13 sources (Fig. 2) are included in the inversion. The source positions were treated as unknown parameters, with starting and estimated values taken to be the measured values, and the source-position uncertainties (Table II) used to define the uncertainties in regularization matrix $\mathbf{H}_1$ of Eq. (4) for the source parameters. The minimum-curvature regularization is applied in x and y but not in z (since vertical smoothness is not physically meaningful for the VLA). To examine AEL results for receiver positions constrained primarily by the acoustic data, prior estimates for inter-receiving spacings are not applied in this case, but are considered subsequently. Also, to avoid using knowledge of the true receiver positions in the inversion, the starting values and prior estimates for the receiver positions were chosen at random from Gaussian distributions centered at the true parameter values with standard deviations of 50 m for x and y and 20 m for z. To obtain error measures that are independent of the starting values, the inversion was repeated 50 times, each time with a different set of random starting values/prior estimates for the receiver positions. Results are considered in terms of the mean error in x, y, z for each receiver (i.e., the absolute difference between the inversion results and the measured receiver positions, averaged over the 50 random initializations) and the standard deviation about these means. These measured errors are compared to the AEL uncertainties estimated from Monte Carlo appraisal with 50 random realizations.

Figure 6 shows the measured and estimated AEL errors, with one standard-deviation error bar included on each. The left column of Fig. 6 shows the inversion results in terms of absolute position errors, i.e., the errors relative to the fixed (geographical) coordinate system. However, for most array processing applications, the relative position errors (i.e., errors in an array-based coordinate system) are a more relevant measure of uncertainty. This is because position errors common to all receivers are equivalent to a simple translation and/or rotation of the receiver array, while relative position errors introduce inter-receiver timing and phase errors which can degrade or preclude some applications such as source localization or environmental inversion. The right column of Fig. 6 shows the AEL results in terms of relative positioning errors, computed as described in Sec. II. The measured errors (absolute and relative) for the receiver positions are also given in Table I.

Figure 6 shows that the Monte Carlo error estimates are similar to, but usually somewhat smaller than, the measured errors. The relative position errors are generally smaller than the absolute errors and have smaller variability. The measured relative errors averaged over all receivers are 13 cm in the vertical, 27 cm in the horizontal (i.e., 2D errors consisting of the square root of sum of squares of x and y errors), and 30 cm in 3D (the actual AEL errors are likely even smaller, since the measured receiver positions are themselves uncertain by 5–7 cm). These results indicate that regularized AEL inversion with imprecise source positions can provide highly accurate receiver positions, and that meaningful uncertainty estimates are provided by Monte Carlo appraisal. It is interesting to note that the 2D horizontal position errors are somewhat smaller for the VLA receivers ($R3-R6$), with

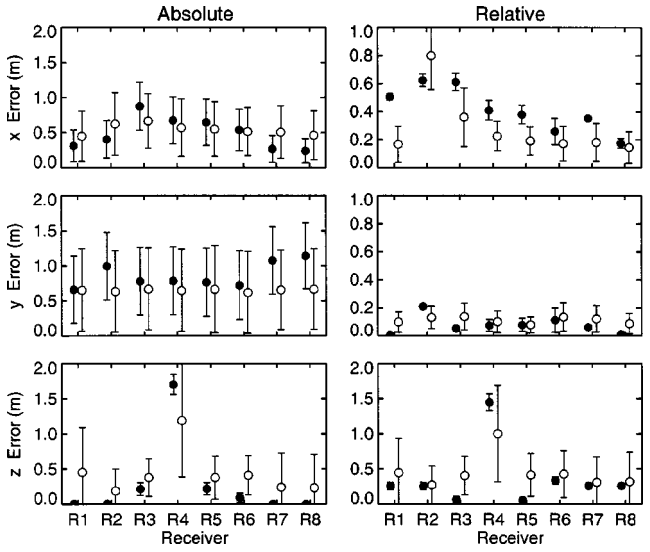


FIG. 7. Receiver localization errors from comparison with measured positions (closed circles) and Monte Carlo analysis (open circles) using data from only five sources, with prior source location uncertainties of 1 m in x and y and 0.5 m in z.

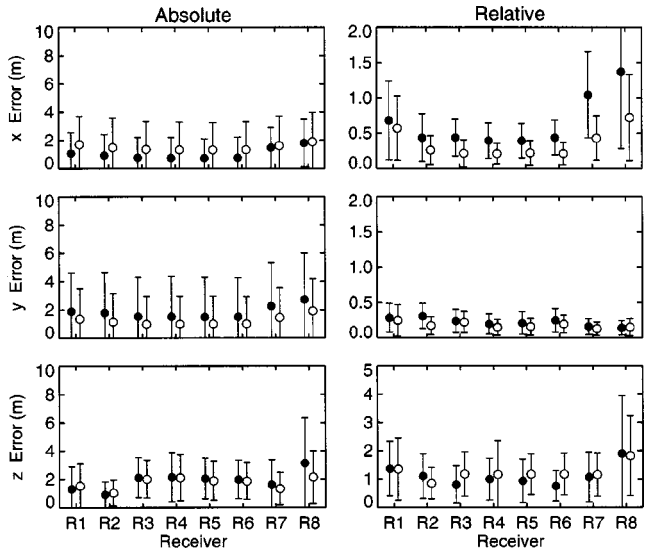


FIG. 8. Receiver localization errors from comparison with measured positions (closed circles) and Monte Carlo analysis (open circles) with prior source location uncertainties of 30 m in x and y and 4 m in z.

an average relative error of 22 cm. This is likely because the distribution of the close-range sources is centered on the VLA (Fig. 2), providing superior horizontal control. Further, the x component of the relative errors is generally larger than the y errors in all cases considered in this paper, likely because since the array is aligned along the x axis, the rotation correction has little effect on these errors (but a significant effect on y errors). Finally, in all inversion results in this paper, the sound-speed bias was found to be negligibly small ($\sim$0.1 m/s).

The AEL inversion results in Fig. 6 made use of all 13 source positions. To investigate the dependence on the number of sources, Fig. 7 shows inversion results computed for an identical problem, except that only five sources were employed: $S2$, $S3$, $S8$, $S9$, and $S13$ (see Fig. 2). Figure 7 shows that good localization results are still obtained, although the mean positioning errors and variability are generally somewhat larger for the smaller number of sources (cf. Fig. 6). The large positioning error in z for VLA receiver $R4$ is due to the large uncertainty associated with the arrival time for the shallow source $S13$ (directly above), as shown in Fig. 5. This error is larger in Fig. 7 than in Fig. 6 since no other short-range sources were included which provide strong vertical-positioning information.

In most at-sea AEL surveys, the uncertainties in source positions are larger than those obtained here for the through-the-ice survey. To consider such cases, random errors were added to the source positions used in the inversion of the measured AEL data (all 13 sources are included here). These errors were drawn from zero-mean Gaussian distributions with standard deviations of 30 m for x and y, and 4 m for z (representative of uncertainties due to GPS positioning and oceanic swells). Similar source-position errors were included in the Monte Carlo appraisal. The results of the AEL inversion with increased source-position errors, shown in Fig. 8, are significantly poorer than those in Fig. 6. However, useful receiver positions are still obtained. The measured relative (2D) horizontal errors are typically are $\lesssim$1 m, with errors of 20–40 cm for the receivers near the center of the source distribution ($R2-R6$). The relative errors in z vary from approximately 1 to 2 m. One way to improve AEL results is to include additional *a priori* information (if available). As an example, Fig. 9 shows AEL results for a case identical to that of Fig. 8, except that prior information of the vertical inter-receiver spacing was included. Figure 9 show a substantial improvement in receiver positions in all three dimensions, with measured relative errors $\lesssim$10 cm in z and 10–60 cm in the horizontal.

To quantify the benefit of including the source positions as unknown parameters in the regularized inversion, Fig. 10 shows AEL results for a case identical to that of Fig. 8 (i.e., 13 sources with large position errors), except that the inac-

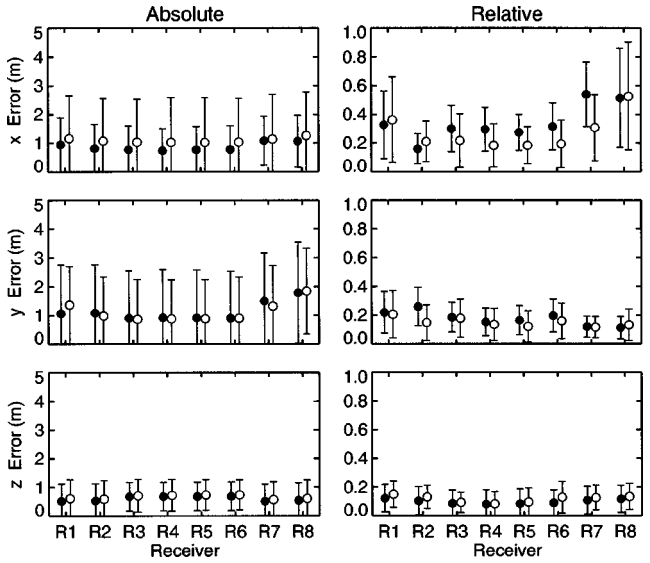


FIG. 9. Receiver localization errors from comparison with measured positions (closed circles) and Monte Carlo analysis (open circles) with prior source location uncertainties of 30 m in x and y and 4 m in z and prior information on vertical sensor spacing.

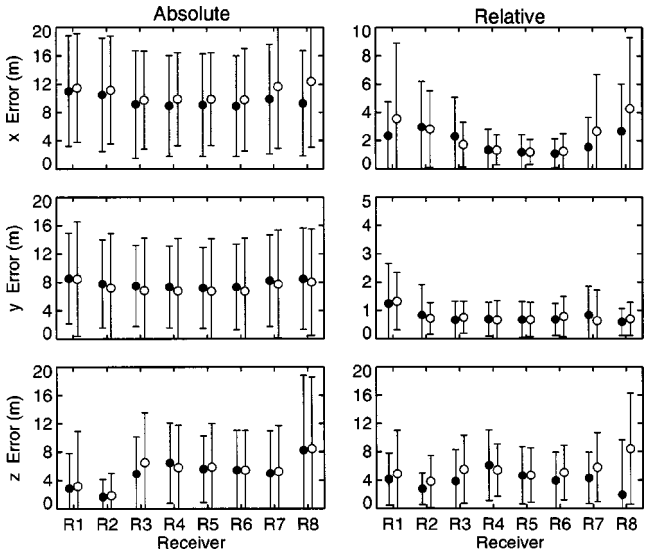


FIG. 10. Receiver localization errors from comparison with measured positions (closed circles) and Monte Carlo analysis (open circles) with prior source location uncertainties of 30 m in x and y and 4 m in z when source locations are not included in the AEL inversion.

curate source positions were held fixed, as is often done in AEL. The results in Fig. 10 are substantially worse than those in Fig. 8, with measured relative errors of 1–3 m in x and y and 2–6 m in z. Even poorer results would be expected for fewer sources. Finally, it should be noted that Figs. 6–10 all show good general agreement between the measured receiver-position errors and the Monte Carlo uncertainty estimates.

V. SUMMARY AND DISCUSSION

Array element localization is an important prerequisite for advanced array processing, but little work has been reported to date that quantitatively examines the accuracy of AEL results. This paper considered the method of regularized AEL by comparing inversion results to direct measurements of receiver positions for a full-scale AEL survey. To achieve this, a stabilized 2D array was deployed from shorefast (motionless) sea ice. Receiver positions were measured to ~5 cm accuracy in each dimension (9 cm in 3D) using nonacoustic methods (laser survey, ROV video). Impulsive sources were deployed around the array, with source positions known to ~1 m. The average measured AEL errors (difference between measured receiver positions and inversion results) were 13 cm in depth, 27 cm in the horizontal, and 30 cm in 3D. Generally good agreement was achieved between the measured errors and AEL uncertainties estimated by Monte Carlo appraisal. These results provide direct validation of the regularized AEL inversion methodology and its underlying assumptions (e.g., Gaussian-distributed data and prior-estimate uncertainties), and represent an important benchmark for acoustic AEL methods.

The effect on receiver-position errors of the number of sources included in the inversion and of the uncertainty in source positions was quantitatively investigated. Good AEL results were obtained with just five sources (four is the minimum possible for AEL with relative travel times). However, AEL results degraded significantly with increasing source-position uncertainties. A substantial increase in receiver-position errors occurred when source positions were not treated as unknown parameters in the inversion, indicating this is an important component of AEL inversion.

The results of this study indicate that even higher precision AEL results could be achieved by more tightly constraining source positions or by including prior information on inter-receiver spacings (if reliable information is available). Higher-precision results can also be obtained by inverting absolute travel-time measurements.[17,18] However, this requires synchronization in time between source transmissions and the recording system, resulting in a more complex AEL survey.

[1] W. S. Burdic, *Underwater Acoustic System Analysis* (Prentice–Hall, Englewood Cliffs, NJ, 1994).
[2] A. Tolstoy, *Matched Field Processing for Underwater Sound* (World Scientific, Singapore, 1993).
[3] A. B. Baggeroer, W. A. Kuperman, and P. N. Mickhalevsky, "An overview of matched field methods in ocean acoustics," IEEE J. Ocean. Eng. **18**, 401–424 (1993).
[4] M. D. Collins, W. A. Kuperman, and H. Schmidt, "Nonlinear inversion for ocean-bottom properties," J. Acoust. Soc. Am. **92**, 2770–2783 (1992).
[5] S. E. Dosso, M. L. Yeremy, J. M. Ozard, and N. R. Chapman, "Estimation of ocean-bottom properties by matched-field inversion of acoustic field data," IEEE J. Ocean. Eng. **18**, 232–239 (1993).
[6] B. D. Steinberg, *Principles of Aperture and Array Systems Design* (Wiley, New York, 1976).
[7] W. S. Hodgkiss, D. E. Ensberg, J. J. Murray, G. L. D'Spain, N. O. Booth, and P. W. Schey, "Direct measurement and matched-field approaches to array shape estimation," IEEE J. Ocean. Eng. **21**, 393–401 (1996).
[8] G. J. Heard, M. McDonald, N. R. Chapman, and L. Jaschke, "Underwater light bulb implosions: A useful acoustic source," Proceedings MTS/IEEE Oceans 97, 1997, Vol. 2, pp. 755–762.
[9] T. C. Yang and C. W. Votaw, "Under ice reflections at frequencies below 1 kHz," J. Acoust. Soc. Am. **70**, 841–851 (1981).
[10] E. C. van Ballegooijen, G. W. M. van Mierlo, C. van Schoonveld, P. M. van der Zalm, A. T. Parsons, and N. H. Field, "Measurement of toward array position, shape, and attitude," IEEE J. Ocean. Eng. **14**, 375–383 (1989).
[11] R. C. Shockley, J. A. Rice, and P. Hursky, "Element localization for bottomed arrays without transponders," J. Acoust. Soc. Am. **95**, 2809 (1994).
[12] M. V. Greening, "Array element localization of rapidly deployable systems," Can. Acoust. **28**, 7–13 (2000).
[13] K. C. Creager and L. M. Dorman, "Location of instruments on the seafloor by joint adjustment of instrument and ship positions," J. Geophys. Res. B **87**, 8379–8388 (1982).
[14] S. E. Dosso, M. R. Fallat, B. J. Sotirin, and J. L. Newton, "Array element localization for horizontal arrays via Occam's inversion," J. Acoust. Soc. Am. **104**, 846–859 (1998).
[15] S. E. Dosso and N. E. Collison, "Regularized inversion for towed-array shape estimation," *Inverse Problems in Underwater Acoustics*, edited by M. Taroudakis and G. Makrakis (Springer, Berlin, 2001).
[16] S. E. Dosso and N. E. B. Collison, "Acoustic tracking of a freely drifting sonobuoy field," J. Acoust. Soc. Am. **111**, 2166–2177 (2002).
[17] S. E. Dosso and B. J. Sotirin, "Optimal array element localization," J. Acoust. Soc. Am. **106**, 3445–3459 (2000).
[18] S. E. Dosso, G. H. Brooke, S. J. Kilistoff, B. J. Sotirin, V. K. McDonald, M. R. Fallat, and N. E. Collison, "High-precision array element localization of vertical line arrays in the Arctic Ocean," IEEE J. Ocean. Eng. **23**, 365–379 (1998).
[19] S. E. Dosso and M. Riedel, "Array element localization for towed marine seismic arrays," J. Acoust. Soc. Am. **110**, 955–966 (2001).
[20] R. I. Verrall, "Tetrahedral underwater arrays and the effect of current on their shape," DREA Tech. Memo. 2000-093, 2000, 16 pp.
[21] R. I. Verrall and G. J. Heard, "Technique for the accurate positioning of objects in an ice-covered ocean," DREA Tech. Memo. 2000-103, 2000, 25 pp.

[22] W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory* (Academic, Orlando, 1984).
[23] C. van Schooneveld, ''Inverse problems: A tutorial survey,'' in *Underwater Acoustic Data Processing*, edited by Y. T. Chan (Kluwer, Dordrecht, 1989), pp. 393–411.
[24] W. H., Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
[25] M. V. Greening and P. Zakarauskas, ''Spatial and source level distributions of ice cracking in the Arctic Ocean,'' J. Acoust. Soc. Am. **95**, 783–790 (1994).
[26] M. V. Greening and P. Zakarauskas, ''Pressure ridging spectrum level and a proposed origin of the infrasonic peak in Arctic ambient noise spectra,'' J. Acoust. Soc. Am. **95**, 791–797 (1994).
[27] S. E. Dosso, G. J. Heard, and M. Vinnins, ''Source bearing estimation in the Arctic using ice-mounted geophones,'' J. Acoust. Soc. Am. **112**, 1390–1398 (2002).

Development of f_2/f_1 ratio functions in humans[a]

Barbara A. Vento,[b] John D. Durrant,[c] and Diane L. Sabo
Department of Communication Science & Disorders, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

J. Robert Boston
Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

(Received 2 September 2003; revised 23 January 2004; accepted 25 January 2004)

Otoacoustic emissions (OAEs) presumably represent active processes within the cochlea fundamental to frequency-selectivity in peripheral auditory function. Maturation of the cochlear amplifier, vis̀-a-vis̀ frequency encoding or selectivity, has yet to be fully characterized in humans. The purpose of this study was to further investigate the maturation of features of the f_2/f_1 frequency ratio (Distortion Product OAE amplitude *X* f_2/f_1 ratio) presumed to reflect cochlear frequency selectivity. A cross-sectional, multivariate study was completed comparing three age groups: pre-term infants, term infants and young adult subjects. Frequency ratio functions were analyzed at three f_2 frequencies—2000, 4000 and 6000 Hz. An analysis included an estimation of the optimal ratio (OR) and a bandwidth-like measure (Q_3). Analysis revealed significant interactions of age *x* frequency *x* gender for optimal ratio and a significant interaction of age *x* frequency for Q_3. Consistent and statistically significant differences for both OR and Q_3 were found in female subjects and when $f_2=2$ or 6 kHz. This supports research by others [Abdala, J. Acoust. Soc. Am. **114**, 3239–3250 (2003)] suggesting that the development of cochlear active mechanisms may still be somewhat in flux at least through term birth Furthermore, OAEs appear to demonstrate gender differences in the course of such maturational changes. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1675819]

PACS numbers: 43.64.Jb [BLM] Pages: 2138–2147

I. INTRODUCTION

The embryologic development of the cochlea begins with formation of the otocyst in the third to sixth week of fetal development. The 2 $\frac{1}{2}$ turns of the cochlea are present at 11 weeks gestational age (GA). Structural development of the cells of the organ of Corti, including the inner and outer hair cells, is complete by 20 weeks GA. The inner hair cells mature first, followed by the outer hairs cells in a base-to-apex gradient, and afferent innervation precedes efferent innervation. At 10–12 weeks GA, both inner and outer hair cells are innervated by afferent fibers only. It is only at the 14th week of gestation that efferent fibers begin to appear. The cochlea and basilar membrane structures and innervation appear mature at 30 weeks GA (Rubel, 1985; Pujol & Lavigne-Rebillard, 1995; Pujol & Lavigne-Rebillard, 1992) as visualized via electron microscopy. Human fetuses have shown responses to sound, in utero, as early as 22 weeks GA (Birnholz and Benacerraf, 1983). However, the presence of adult-like structures in the cochlea and the onset of function does not necessarily equate with mature auditory function (Pujol & Lavigne-Rebillard, 1992). The time course for the development of frequency coding in human infants is unknown but was generally accepted to be largely adult-like by term birth (Abdala *et al.*, 1996; Abdala and Sininger, 1996; Lasky, 1998a; Lasky, 1998b). However, results of recent investigations suggest that, in fact, ''subtle immaturities'' in cochlear function appear to continue, even after term birth (Abdala and Chatterjee, 2003; Abdala and Fitzgerald, 2003; Abdala, 2003). Further investigation of this developing time course of frequency coding would be of scientific and potentially clinical interest. Otoacoustic emission measurement provide a method of investigating cochlear function, even in very young pre-term neonates.

Researchers have completed several studies investigating the developing infant cochlea with otoacoustic emissions (Smurzynski, 1994; Brown *et al.*, 1995; Abdala, 1996; Morlet *et al.*, 1996, 2003; Abdala, 1998; Lasky, 1998a, b; Abdala, 2000, 2001a, 2003). The absolute amplitude of both transient and distortion product otoacoustic emissions are known to increase throughout infancy (Smurzynski, 1994). However, there is some question as to whether the increases are the same in both sexes and similar for all frequencies (Morlet *et al.*, 1996). Others have utilized suppression tuning curves (STC) of the distortion product otoacoustic emission (DPOAE) to follow the frequency maturation of the developing cochlea (Abdala, 1998; Abdala *et al.*, 1999; Abdala, 2001a, b; Abdala, 2003). Although the STCs showed only very subtle non-adult-like characteristics and were basically adult-like in term born infants, they differed in pre-term infants, especially those younger than 35 weeks GA. This was not the case for DPOAE suppression growth in pre-term infants, at term PCA and full-term neonates (Abdala and Chatterjee, 2003). These infants showed differences from adults, especially at $f_2=6$ kHz.

Brown and colleagues (1993) completed a study com-

[a]Portions presented at Biennial Conference of the International Evoked Audiometric Response Study Group, Vancouver, BC, July 2001.
[b]Corresponding author; electronic mail barbv@pitt.edu
[c]Also at the Department of Otolaryngology.

paring the f_2/f_1 ratio function (FR, the function of DPOAE magnitude versus f_2/f_1) and psychophysical equivalent rectangular bandwidth (ERB) measurements in adults. The results demonstrated a negative correlation between the peak of the frequency ratio (FR) function and the ERB at f_2. Yet, this relationship allows an objective and noninvasive means of studying frequency selectivity in the cochlea using the FR functions that is less time intensive than the psychophysical measurements. The FR function, consequently, has been employed to investigate cochlear frequency maturation in human pre-term infants in some studies (Abdala, 1996; Lasky, 1998a, b; Brown and Gaskil, 1990; Brown *et al.*, 1993, 1995). These investigators reported differences among age groups but the differences seldom reached statistical significance. This may have been due to the method of determining the optimal ratio (OR) of the function and/or the test parameters used. In addition, although we know of gender differences in auditory function, these studies did not control for this variable.

Although FR functions are not a direct measure of frequency tuning, they logically reflect common underlying micro-mechanical events based on the widely held concept that OAEs are manifestations of cochlear active processes. The measurements involved are relatively efficient to administer, even in the premature human neonate. They thus provide an objective and noninvasive means of investigating the maturation of cochlear active mechanisms that are presumably requisite to frequency coding and/or resolution in the cochlea. Indeed, previous investigators have used FR functions to study cochlear maturation in neonates (Brown *et al.*, 1994, 1995; Abdala, 1996; Lasky, 1998a; Brown *et al.*, 2000). The present study was, in part, a replication of this earlier work but was dedicated to further investigation of features of the FR function in premature neonates versus term neonates and young adults. Of particular interest were the parameters of optimal ratio and selectivity (Q_3, as defined below), and the analysis of the data by age while controlling for gender, a factor that has only recently begun to receive attention (Morlet *et al.*, 1996, 2003).

II. METHODS

A. Subjects

This study was approved for the use of human subjects by the Internal Review Board (IRB) of the University of Pittsburgh. All subjects or their parent/guardian completed informed consent prior to testing.

Thirty weeks GA is the time the cochlea is believed to be structurally, if not functionally, mature (Rubel, 1985; Pujol & Lavigne-Rebillard, 1992, 1995). Therefore, the target age range of the pre-term infants was chosen to be 30–33 weeks post conceptional age (PCA). The second group comprised term infants for comparison of auditory function at full term birth. Finally, a young adult group (18–25 years) was examined, as controls, for a comparison to the widely accepted standard of mature auditory system function.

Pre-term infants were identified by chart review at Magee Women's Hospital (MWH) of the University of Pittsburgh Medical Center. The PCA was determined by adding the chronological age (CA) to the gestational age of the infant as reported in the medical record. To effectively assess normal auditory development, it was important that the young pre-term infants be healthy and as free from potential environmental effects as possible. Premature subjects received a prophylactic standard antibiotic course in the NICU to rule out sepsis. Other than this, they had no other high risk factors for hearing loss. A total of 23 pre-term infants, 11 male subjects with a mean PCA of 31.1 weeks and 12 female subjects with a mean PCA of 31.2 weeks, were successfully recruited and tested.

Term infants were recruited from the regular care nurseries of MWH. All full-term newborns 37–42 weeks GA and meeting all pre-term criteria were eligible for inclusion in the subject pool. Thirty-six term infants, 18 male with a mean GA of 39.2 and 18 female mean GA of 39.6 subjects were tested according to the protocol.

The infants were required to pass a DPOAE screening test according to standards set for newborn infant hearing screening at MWH. Specifically, for inclusion in this study, a pass was defined as a response of at least 8 dB SPL above the noise floor for at least four of the f_2 frequencies tested. Stimulus parameters were L1 = 65 dB SPL, L2 = 55 dB SPL with $f_2/f_1 = 1.2$ for six f_2 frequencies from 2.5 kHz to 8 kHz.

Adult subjects were recruited from the student population of the University of Pittsburgh and the Pittsburgh community. The mean age for female subjects was 22.6, and for males, 21.7. Adults had audiometric thresholds <15 dB HL from 500 to 8000 Hz and a negative history of otologic disease. None of the subjects had high risk factors for hearing loss. There were 36 adult subjects, 18 Caucasian women, 11 Caucasian, and 7 Hispanic males, between 18 and 25 years of age who met subject selection criteria and were subsequently tested.

Ninety-five subjects, in total, completed the study protocol. However, due to incomplete data sets, only 85 subjects were included in the frequency ratio experiment at the 65/55 stimulus level. There were 18 pre-term infants (9 male, 9 female), 32 term infants (15 male, 17 female), and 35 adults (18 male, 17 female). Subjects were eliminated from the analysis if they did not, in effect, complete all three frequencies (waking before the completion of testing), if there were insufficient data points in a particular FR function [fewer than five data points with a signal to noise ratio (SNR) of at least 3 dB], or if they had bi-peak functions, as defined below.

A total of 73 subjects were included in the analysis of Q_3 at the 65/55 stimulus level. There were 16 pre-term infants (8 female, 8 male), 28 term infants (14 female, 14 male) and 29 adults (14 female and 15 males). Twelve subjects were not included due to FR functions that did not conform to the typical shape. Specifically, these functions did not "peak" sufficiently to allow for calculation of the f_2/f_1 ratio width, vis-à-vis criteria adopted for the determination of Q_3. Function width was determined at the 3-dB-down points referenced to the peak of the function identified by the OR. This measurement is illustrated in Fig. 1.

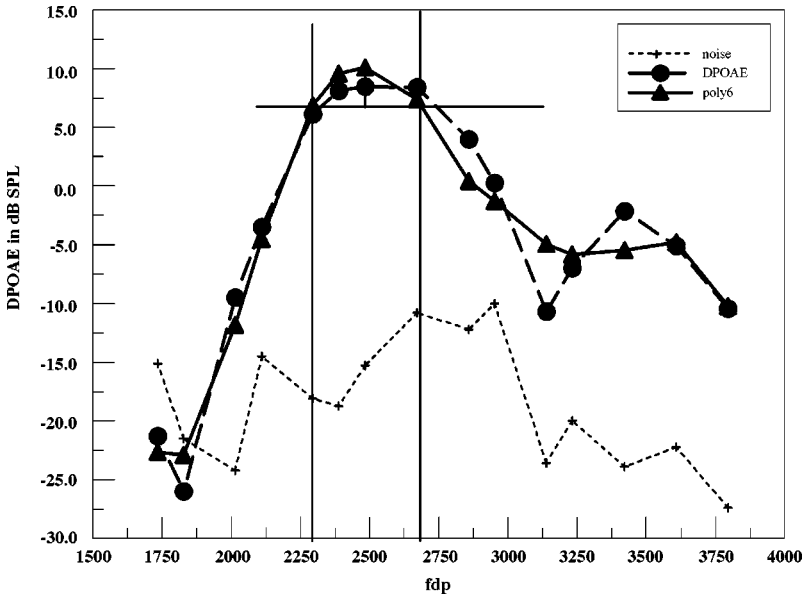


FIG. 1. Example of a $f2/f1$ ratio function. The dashed line represents the raw data; the solid line represents the 6th order polynomial fit of the data. The horizontal and vertical lines denote the 3 dB down point for the Q_3 calculation.

B. Equipment

The Etymotic Research ER-10C OAE probe microphone was used to monitor DPOAEs in conjunction with the Scout DPOAE program from Bio-logic Systems Corporation. The ER-10C has independent transducers for the two primary tones. Primary levels were maintained at the desired levels according to ear canal calibration measured by the microphone in the probe. Noise floor estimations were computed from the average noise in the frequency bins on either side of the frequency of the distortion product $2f_1$-f_2; the bins were 50 Hz in width. Averaging continued at each FR until the noise was equal to or less than -15 dB SPL, the signal to noise (S/N) ratio was at least 8 dB SPL, or the ''point time'' limit (i.e., 10 seconds/data point) was reached. Specific f_2/f_1 ratios were between 1.01 through 1.40 for each of the f_2 frequencies tested. The f_1 frequencies for a given f_2 frequency, were computed and controlled by the Scout program according to the range of interest (above). The number of ratios for each function ranged from 13–16 depending on f_2 frequency. System distortion was tested in the Zwislocki coupler and found to be -15.8 dB SPL or less at f_2 frequencies between 1453 and 10,028 Hz, which was considered the practical frequency range for purposes of this study.

Hearing threshold testing for the adult subjects was completed in a double-walled sound-treated room (Industrial Acoustics Company, Inc.) using a Beltone 2000 Clinical Audiometer. Subjects passing the hearing screening completed all remaining experimental conditions. The audiometer was electroacoustically calibrated according to ANSI Standards and biologic checks were conducted before each test session.

C. Procedures

The infants were tested in the regular and neonatal nurseries of Magee Women's Hospital while sleeping. Adult subjects were tested in laboratory space at the University of Pittsburgh. Although, the adults' hearing sensitivity was screened in a sound room, DPOAE testing was completed outside of the sound room. This more closely resembled nursery, background noise conditions. Testing was stopped after 60 minutes for the infants.

The DPOAE was measured from $f2=8000$ Hz to $f2=2500$ Hz (i.e., the so-called, ''DP-gram'') at 3 points per octave for a total of six frequencies in both ears for all subjects. A ''pass'' on this DP-gram was taken to indicate normal peripheral auditory function and the absence of middle ear dysfunction (Chang *et al.* 1993; Trine *et al.*, 1993; Owens *et al.*, 1993; Wada *et al.*, 1995).

Frequency ratio functions were analyzed to estimate optimal f_2/f_1 ratios (OR). The FR ratio is obtained by varying the relationship of the primary tones, f_2/f_1, where f_2 is held constant and the f_1 frequency is swept between ratios of 1.01 to 1.40. The Scout program selected f_1 frequencies for every f_2 frequency tested. The f_2 frequencies tested were 2 kHz, 4 kHz and 6 kHz. These frequencies were selected in an effort to assess a reasonable area of frequency coding along the basilar membrane. Given difficulties in obtaining reliable information in lower frequencies in pre-term and term infants, $f_{2=}2$ kHz was selected as the lowest best frequency. Other researchers (Brown *et al.* 1995, Lasky, 1998a) have shown differences between infants and adults at $f_{2=}4$ kHz that were not noted at other frequencies on FR function tests. Therefore, $f_{2=}4$ kHz was chosen as a second frequency. Finally, $f_{2=}6$ kHz FR functions were reported to be very similar across age groups (Lasky, 1998a; Abdala, 1996). This f_2 was chosen to see if the similarities remained when gender and ear were included individually as variables. At $f_2=2$ kHz, f_1 was swept from 1990 Hz ($f_2/f_1=1.01$) to 1430 Hz ($f_2/f_1=1.40$). At $f_2=4$ kHz, f_1 was swept from 3950 Hz ($f_2/f_1=1.01$) to 2950 Hz ($f_2/f_1=1.37$). At $f_2=6$ kHz, f_1 was swept from 5940 Hz ($f_2/f_1=1.01$) to 4290 Hz ($f_2/f_1=1.40$). There were two levels of L1 (L1–L2=10 dB) measured for each f_2 frequency, 65/55 and 50/40 (L1/L2) dB SPL. The order of testing of the FR functions was randomized by frequency. Measurements at the 65/55 dB intensity level was completed first in an attempt to obtain complete set of data for the infant subjects for at least one intensity level,

TABLE I. Means and standard deviations for optimal ratios.

Parameters /groups		2 kHz		4 kHz		6 kHz	
		Mean	s.d.	Mean	s.d.	Mean	s.d.
Pre-term	F	1.234	0.0380	1.214	0.0315	1.152	0.0293
	M	1.241	0.0189	1.221	0.0415	1.197	0.0367
Term	F	1.205	0.0345	1.195	0.0413	1.175	0.0256
	M	1.223	0.0468	1.225	0.0327	1.193	0.0277
Adult	F	1.221	0.0359	1.232	0.0182	1.222	0.0557
	M	1.248	0.0305	1.245	0.0289	1.203	0.0470

namely the level most likely (*a priori*) to yield adequate signal-to-noise ratios for all frequencies of interest.

A 6th order polynomial lest-squares fit was applied to all data points for each subject and condition (Fig. 1). Each subject was required to have five data points greater than 3 dB above the noise floor in the FR function to allow for determination of OR. After the FR function was fitted, the derivative was calculated to identify the point where the function peaked. For a few cases, the derivative identified two peaks. Inspection of the graphs readily identified such functions as bi-peaked, even before fitting. By definition, the determination of OR and the analysis of selectivity require a single-point estimate of OR. Therefore, for purposes of this study, bi-peaked FR functions were considered anomalous and were not included in the optimal-ratio analysis. Although conceptually worrisome at the inception of the study (vis-à-vis use of a 6th order fit), the exclusion of data sets according to this criterion proved rare (i.e., 3 subjects: 1 male pre-term, 1 female pre-term, and 1 male term infant.)

Finally, selectivity or Q of the frequency ratio function was evaluated. It was necessary to first determine function width (FW), namely measuring the bandwidth at the point 3-dB-down from the peak of the function or optimal ratio (Fig. 1). The optimal ratio was used to calculate the frequency of the distortion product (fdp) represented by the optimal ratio for a given function ($f_2/\mathrm{OR}=f_1$). The fdp was calculated, naturally, as $2f1\text{-}f2$ Likewise, the ratios for the width measurements were converted to fdp. The higher ratio (at -3 dB) was designated fdp_a and the lower ratio fdp_b. The width was calculated as $\mathrm{fdp}_b-\mathrm{fdp}_a=\mathrm{fdp}_w$. Finally, Q_3 was calculated as $\mathrm{fdp}_{or}/\mathrm{fdp}_w$.

D. Statistical analysis

This study was designed to further explore the maturation of the cochlear amplifier (as manifested in the FR function and Q_3) and to test if gender impacts the interpretation of DPOAE measurements in the tracking of development of the cochlear amplifier in humans. A multifactorial analysis of variance (MANOVA) was conducted. Optimal ratios were submitted to a three-way ANOVA. Interactions were considered significant at $p=0.05$. *Post hoc* analysis was completed using the simple–simple main or simple main effects tests. A *post hoc* comparison was considered significant at $\alpha=0.05$.

III. RESULTS

A. Optimal ratios

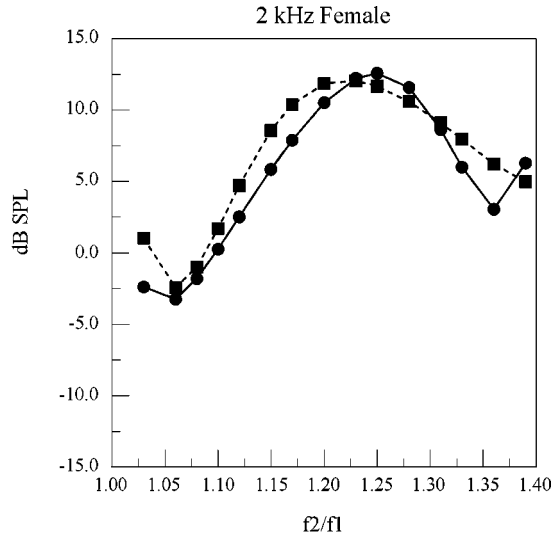
Optimal ratios measured under the 65/55 level condition showed a significant three-way interaction of age *x* frequency *x* gender ($\mathrm{df}=4$; $f=3.06$; $p=.012$). Means and standard deviations for the optimal ratios are found in Table I.

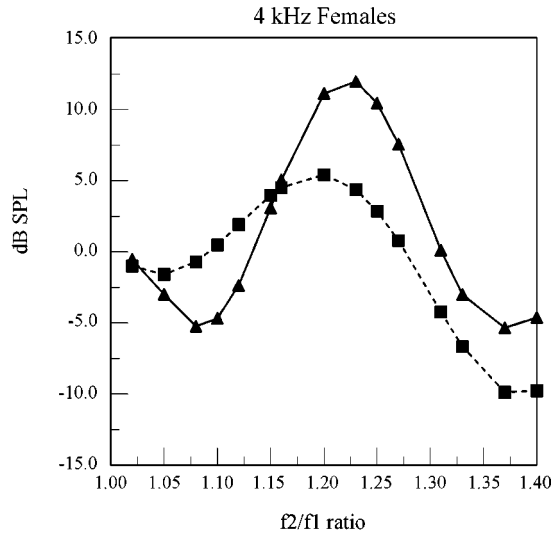
Optimal ratios for males were similar across age groups. On the other hand, statistically significant differences were found among between age groups for the female subjects. Specifically, there was a significant difference observed at 2 kHz between pre-term and term females ($t_{(79)}=2.04$, $p<0.05$). At 4 kHz, differences were significant between term and adult female participants ($t_{(79)}=2.98$, $p<0.01$). Finally, at 6 kHz, differences were significant between both the pre-term and term female infants ($t_{(79)}=4.93$ and 3.79, $p<0.01$) and the adult female subjects. Figure 2 graphically displays the grand averages of the FR functions for female subjects.

Additional evaluation of ORs revealed differences between male and female pre-term infants that were statistically different at 6 kHz ($t_{(79)}=2.647$, $p<0.01$). Gender differences also were statistically significant in term infants at 4 kHz ($t_{(79)}=2.344$, $p<0.05$) and in adults subjects at 2 kHz ($t_{(79)}=2.24$, $p<0.05$). Grand averages for the significantly different FR functions for gender are shown in Fig. 3

Optimal ratios were compared further by frequency and age group for males versus females (Fig. 4). Results showed that optimal ratios for $f_2=2$ kHz and 4 kHz were similar across age groups and gender. However, ORs at 6 kHz were significantly different from those determined at 2 and 4 kHz that were age- and gender-specific as described below.

In male subjects, ORs for $f_2=6$ kHz tended to be significantly different from the ORs of both lower frequency f_2 (Term 2/4 kHz $=t_{(158)}=2.57$, $p<0.05$ and $t_{(158)}=2.74$, $p<0.01$; adults 2/4 kHz ($t_{(158)}=4.22$ and 3.94, $p<0.01$). In pre-term male infants differences were significant between 2 kHz and 6 kHz only ($t_{(158)}=2.92$, $p<0.01$). Female subjects demonstrated a different pattern of change. Specifically, ORs were similar for female adult subjects for all $f2$ frequencies. Statistically significant differences in ORs for the female subjects were found between 2 and 6 kHz in term infants ($t_{(158)}=2.73$, $p<0.01$). In the pre-term female subjects, ORs for 2 and 4 kHz were similar, but both frequencies demonstrated significantly different ORs from 6 kHz (2 kHz, $t_{(158)}=5.44$, $p<0.01$; 4 kHz, $t_{(158)}=4.11$, $p<0.01$).





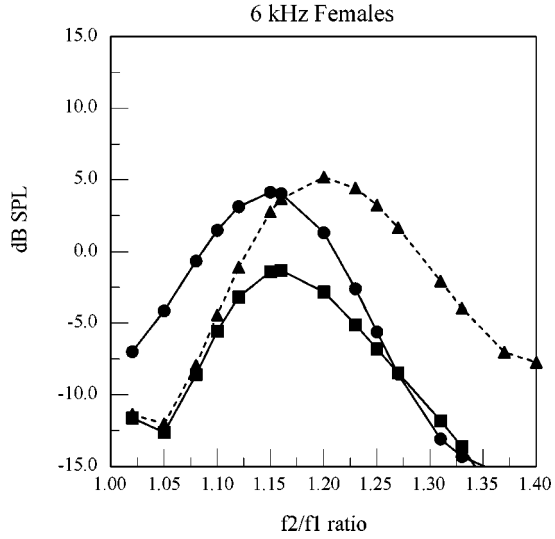
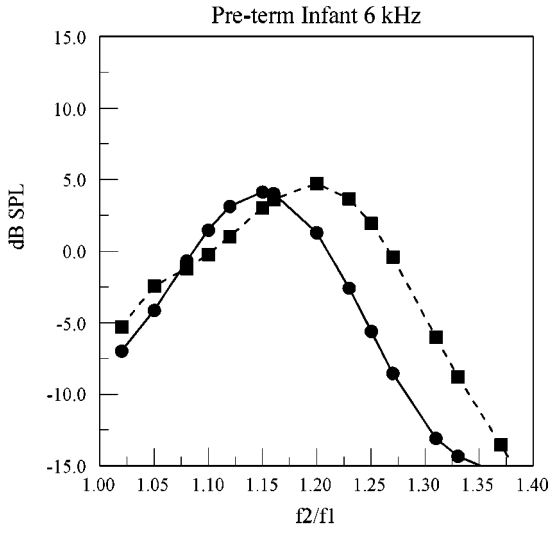
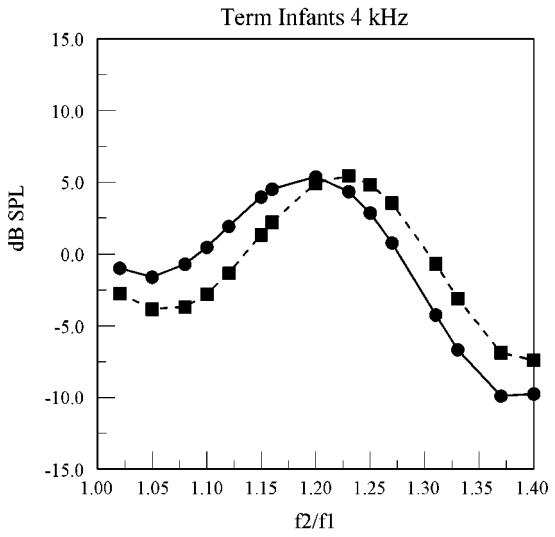


FIG. 2. Pre-term: ●; term: ■; adult: ▲ These figures show the $f2/f1$ ratio grand averages for female subjects separated by frequency. Only statistically significant differences are provided.





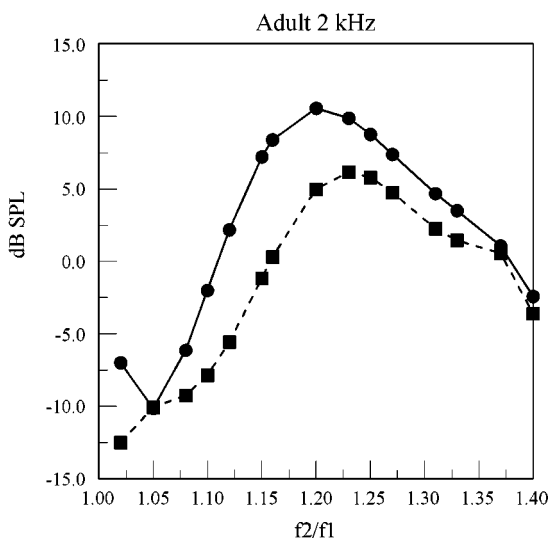


FIG. 3. These are grand averages of the $f2/f1$ ratio functions showing the statistically significant gender X frequency differences for all age groups. Female subjects are represented by circles and male subjects by squares.

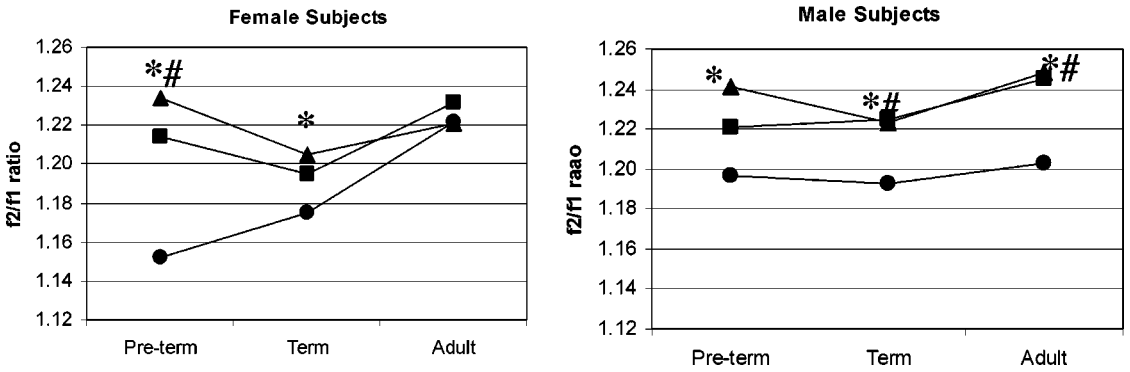


FIG. 4. 2 kHz: ▲; 4 kHz: ■; kHz ●. These graphs display the interactions of the frequency by age group of optimal ratios for males and females. Significant differences are indicated by an* between 2 and 6 kHz and the # for differences between 4 and 6 kHz.

B. Q_3

The results of the ANOVA of Q_3 showed a significant age *x* frequency interaction (df=2; $f=4.371$; $p=0.019$). Means and standard deviations for Q_3 values are found in Table II, and age *x* frequency interactions are shown in Fig. 5. Pre-term infants showed increasing Q_3 by frequency, but differences across frequency did not reach statistical significance. Term infants demonstrated an increase in Q_3 over frequency that is the highest value being observed at 6 kHz. This value was found to be significantly different from Q_3 at 2 kHz and 4 kHz ($t_{(67)}=3.74$; $p<0.01$; $t_{(67)}=2.66$; $p<0.01$). Interestingly, the adult subjects' data showed a trend toward a decrease in Q_3 from 2 kHz to 6 kHz but differences were not statistically significant. In brief, Q_3 for both infant groups was found to increase with increasing frequency. Adults showed the lowest Q_3 for 6 kHz. Returning to the term infants' data, the pattern of Q_3-by-frequency differences between term infants and adults (forming an "X" pattern in Fig. 5) appeared to be a significant feature in and of itself, judging from the finding of statistically significant differences in Q_3 at both 2 and 6 kHz ($t_{(67)}=2.11$ and 2.43; $p<0.05$, respectively).

C. Results at lower intensity levels

Measurement of FR functions was less successful under the 50/40 dB condition, but an analysis of the data available seemed worthwhile as a preliminary study. Following inclusion guidelines for the FR analysis at 65/55 dB, complete data sets were available at all three frequencies for only five pre-term infants. This was largely due to lack of results for $f_2=2$ kHz. With $f_2=2$ kHz excluded, the total number of pre-term infants with complete FR functions increased to eight, with concomitant increases in the number of subjects in the term and adult groups. The final subject total was 9 pre-term infants (5 females, 4 males), 18 term infants (11 females, 7 males) and 35 adults (18 females and 17 males). A three-way ANOVA was completed with two f_2 frequencies, 4 kHz and 6 kHz.

The results showed a significant age *x* frequency *x* gender interaction (df=2; $f=6.99$; $p=.002$). Statistically significant differences were found at $f_2=4$ kHz among pre-term and term female subjects ($t_{60}=3.37$, $p<0.01$) and between term infants and adults ($t_{60}=4.47$, $p<0.01$). There were no significant differences among males. Comparisons of males and females by age group resulted in significant differences at 4 kHz in term infants ($t_{60}=5.54$, $p<0.01$) and at 6 kHz in pre-term infants ($t_{60}=2.06$, $p<0.05$) with males having the higher ratios in both cases. These results are similar to findings at the higher stimulus level (reported above).

There were no statistically significant differences among groups for Q_3 at the lower stimulus level. Inspection of the data (see Fig. 6), including $f_2=2$ kHz, reveals an interesting trend wherein pre-term infants showed the greatest Q_3 values for 2 kHz, a pattern notably different from that observed at 65/55 (described earlier). However, the small sample size did not permit a meaningful statistical analysis.

IV. DISCUSSION

A specific link to frequency resolution of the auditory system and/or causal relationships of frequency encoding or selectivity to FR measures has yet to be proven. Still, frequency ratio functions provide an objective and noninvasive means of exploring the maturation of OAE measures, including premature newborns, thus providing metrics of cochlear active mechanisms believed to underlie frequency encoding in the cochlea. Previous researchers, indeed, have suggested that FR functions indirectly reflect frequency tuning characteristics of the cochlea and that FR measures correlate with psychophysical measures of tuning (Brown *et al.*, 1990,

TABLE II. Means and standard deviations for Q_3.

Parameters /groups	2 kHz		4 kHz		6 kHz	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Pre-term	4.62	1.36	4.71	1.12	5.17	1.63
Term	4.35	1.22	4.72	1.02	5.62	1.13
Adult	5.12	1.50	5.00	1.38	4.73	1.44

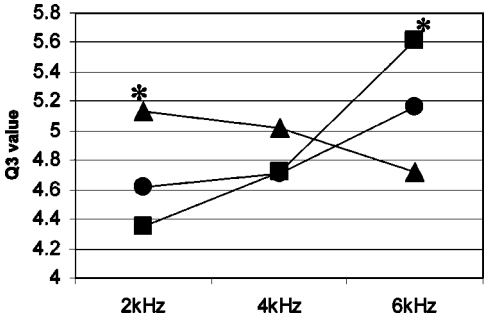


FIG. 5. Pre-term: ●; term: ■; adult: ▲. This graph displays the interaction of Q_3 for age X frequency.* indicates statistically significant differences between the term and adult subjects.

1992, 1993; Kane and Durrant, 1997). On the assumption of some functional significance of the observed changes in FR parameters, the results of the present study support previous findings (Abdala and Chatterjee, 2003; Abdala, 2003), suggesting that maturation of the cochlea is, indeed, incomplete late into the third trimester. These ''late'' changes impact multiple OAE measures—OAE spectrum (Morlet *et al.* 1996, 2003) and OR and Q_3 (present study)—and at multiple frequencies, yet such changes do not follow a straightforward pattern or progression over PCA. It is in this sense that the results are interpreted here as manifestations of cochlear-amplifier-system development which is more or less in a state of flux (for want of a better term). Especially intriguing are the findings that the measures investigated appear not to be adult-like in term, as well as pre-term, neonates and that there are gender differences in the maturational changes observed. The latter suggests that averaging data across the sexes may mask maturational changes reflected in the OAEs. These findings are now discussed in more detail.

A. Optimal ratios

Similar to results of previous studies (Harris *et al.*, 1989; Abdala, 1996; Moulin, 2000), findings of this investigation revealed that the optimal ratio decreases with increasing frequency. In particular, the results reported herein showed that this change occurs most dramatically in pre-term female subjects (Fig. 4). The further observation of results for all female subjects shows that ORs across age groups for $f_2=6$ kHz were the most different (Fig. 2 and 4). Abdala (1996) reported that the optimal ratio, as determined by frequency ratio functions, differed by frequency (1500 and 6000 Hz) but not by age when comparing pre-term, term and adult subjects. She reported that optimal ratio decreases with increasing frequency when collapsed across all other conditions (age and level considered in her study). Moulin (2000) has reported results from an investigation using FR functions in adults. She noted that the optimal ratio decreased with increasing frequency except at 6 kHz where the optimal ratio was significantly larger than the ratios obtained at 4 and 5 kHz. Neither Abdala nor Moulin considered gender as a variable. Given the known structural differences in cochlear length between males and females (Sato *et al.*, 1991), and reports of gender differences in this study and others (Morlet *et al.*, 1996, 2003) research investigations addressing cochlear processing should account for this variable.

The findings of this study differ from those of Moulin's (2000) in that the optimal ratio for 6 kHz in adults did not get smaller. Moulin attributed her observed increase in OR to standing waves, which she supposed would decrease the SPL at the tympanic membrane and lower of intensity primary frequencies, leading to higher f_2/f_1 ratios (Harris *et al.*, 1989). However, standing waves are likely to vary across subjects and are as likely to increase as to decrease in SPL. Abdala (1996) showed smaller optimal ratios at 6 kHz than at 1.5 kHz, as observed in the present study. Abdala used the same probe assembly as was used in the present study, whereas Moulin's had slight methodological differences, including a difference in the stimulus level separation employed. Moulin also had more women in her subject pool than men; this could have increased the average optimal ratio, due to a gender effect. Women in the present study, again, had higher f_2/f_1 ratios at 6 kHz than did the male subjects. Therefore, procedural differences could reasonably account for differences in findings.

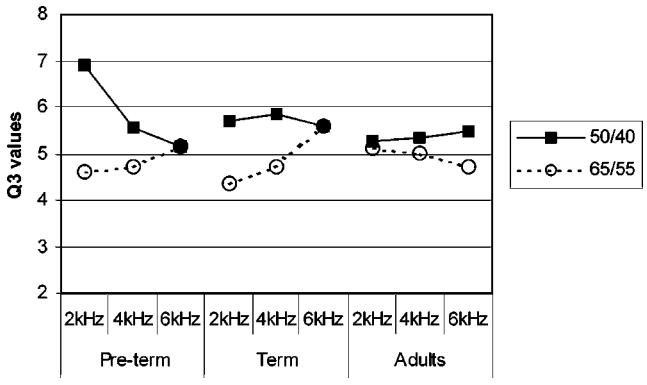


FIG. 6. This graph displays Q_3 means by age, frequency, and level.

B. Q_3 values

In single unit and psychophysical tuning functions, Q values are accepted measures of frequency selectivity. Brown and colleagues (1993) have previously suggested that Q_3 values for FR functions relate to equivalent rectangular bandwidths providing an objective method of measuring frequency selectivity in the cochlea. An analysis of Q_3 values in this study showed a significant frequency *x* age interaction. Results for $f_2=4$ kHz were similar across age groups with statistically significant differences at 2 and 6 kHz between the adult and term infant groups. Pre-term infants showed similar trends to the term infants, but their results were not significantly different from those of adults. Infant Q_3*s* increased with increasing frequency (Fig. 5). In adults, Q_3 became smaller with increasing frequency, implying a decrease in tuning. However, there were no significant differences in Q_3 for the adult subjects in this study. This is different from Q_3 results reported by Abdala (1996). Her results showed increased Q_3 with increasing frequency. She did not report any interactions between age, frequency, or level separation of the primaries (with results collapsed across stimulus levels).

It is noteworthy that for both measures used in the present study differences were found among pre-term infants, term infants, and adults and were present across frequency. The apparent reduction in tuning (lower Q_3 at 6 kHz) for adults raised the question of whether this study's results support early, sub-clinical aging of the adult auditory system or immaturity of the infant auditory system. There is insufficient data from this study to fully support either theory. However, an investigation by Abdala (2001a) specifically addressed this question using DPOAE suppression tuning curves. She tested a group of normal hearing school-aged children with negative histories of potentially adverse environmental exposure. Findings revealed that STCs of the children were similar to the adult subjects and that the DPOAE suppression for the pre-term infants was, in fact, significantly different from both the child and adult groups. Abdala concluded that these findings were more likely evidence of an immature auditory system than premature subclinical aging of the adult subjects. If we accept the interpretation of the findings by Abdala, the results reported in the present study further support the notion of continuing maturational changes in human cochlear function throughout the third trimester, if not beyond.

C. Question of middle ear developmental effects

An implicit concern in OAE measurements in the very young subjects studied is the impact of ongoing maturation of the middle ear system. This is difficult to address completely, especially in pre-term infants, due to limited information on the functional status of their middle ear. The most recent information on infant middle ear function derives from analyses of power reflectance. Keefe and his associates (1993, 1994, 2000) have reported on the development of the middle (and external) ear in humans. In healthy full-term infants tested between 1 to 24 months CA, this maturation is incomplete at 24 months, evidenced by lower power transfer in infants than adults (Keefe *et al.*, 1993). Further investigations, involving pre-term and high risk newborns were completed and showed effects of gender and ear difference in such infants (Keefe *et al.*, 2000). Specifically, the left ear was acoustically stiffer than the right ear. These findings have potential relevance to the interpretation of results from the present study; however, the observed gender differences fell below 2000 Hz, the lowest frequency tested in the present study. Ear differences, on the other hand, were not assessed in this study. It can only be assumed that they were effectively averaged out via the protocol wherein half the subjects were tested in the right and the other in the left, with only one ear being tested (in the interest of economy of time).

There also is the question of how the middle ear maturational changes reported by Keefe and colleagues (2000) might impact measures of cochlear frequency selectivity, *per se*? Pre-term infants, <33 weeks PCA, showed higher reflectance than all other age groups from 33–48 weeks PCA, although changes that were more subtle than those reported previously in older infants and children (Keefe *et al.*, 1993). Reduced power transfer through the middle ear might result in a smaller area of stimulation along the basilar membrane (i.e., sharper tuning) if the test levels are pushing the limits of the cochlear amplifier (as some researchers suspect, in effect, of typical DPOAE stimulus levels (for example see Brown *et al.*, 1994). Middle-ear immaturity might then be implicated in the observation from the present study of a larger Q_3 for pre-term infants (re: term infants) at 2 kHz. Still, this would not explain the reverse situation found at 6 kHz (Fig. 5). Furthermore, the predicted direction of change in tuning is also opposite the change observed in Q_3 at the lower test levels (see below), effectively simulating decreased power transfer. It thus is difficult to attribute the somewhat complex pattern of results from this study to an effective reduction in middle-ear power transfer.

D. Observations at lower levels of stimulation

The higher of the two intensity levels used in this study, 65/55 dB SPL, is often used clinically and is known to produce robust responses in normally functioning cochleas (Whitehead *et al.*, 1995). Testing a lower intensity levels in pre-term infants of <33 weeks PCA is known to be difficult, if not impossible (Abdala, 2003). Abdala reported that, in fact, many infants at 33 weeks PCA do not show DPOAEs at 6 kHz with low level primary tones (40 dB). Therefore, our lack of success at 2 kHz with the added problem of physiologic noise is not surprising. Still, the use of lower stimulus levels is appealing to assure the sampling uniquely of the ''active'' portion of the DPOAE I/O function in infants. The fact that testing at $f_2=6$ kHz was more successful than at $f_2=2$ kHz in the present study is itself an interesting observation, yet one that is not completely in agreement with Abdala's experinence. However, Abdala was utilizing a fixed f_2/f_1 ratio whereas the optimal ratio was the focus here, perhaps facilitating in the process the measurement of responses at 6 kHz.

Despite the limitations of the data set realized, the lower level results permitted the detection of a significant three way interaction for age *x* gender *x* frequency for an optimal ratio. Unlike findings at higher SPLs, which seldom showed differences at 4 kHz, statistically significant differences were observed at 4 kHz (50/40). This could be a reflection of immaturity in CAMs not demonstrated at the higher intensity level.

Although the ANOVA of the Q_3 data for the 50/40 stimulus level did not demonstrate statistically significant effects, a comparison of results for the two stimulus levels shows some interesting patterns, as seen in Fig. 6. There appears to be a systematic change in the pattern of Q_3 in the infant groups at the lower intensity level. Recent work by Abdala (2003) showed that DPOAE detection thresholds at 6 kHz in infants at 33 week PCA is elevated compared to older infants and adults and has a significant variability. Therefore, given the reduction in the number of subjects at 2 kHz at the 50/40 stimulus level, especially for the pre-term infants, the effects reported here must be considered tentative. Additional studies focusing on lower intensity levels (whereas priority here was given to the higher levels), and perhaps lower frequencies (whereas priority here was given to frequencies above 2 kHz), are needed to further substantiate these pre-

liminary findings. At the very least, nevertheless, these results suggest the potential efficacy of working at lower levels than commonly used, with persistence.

E. Clinical application and future study

One clinical implication that may be derived from the results of this study is the possible value of the use of optimal frequency ratios specific to age groups. First, it is important to note that the findings reported here do not mitigate, for purposes of a universal test, the choice of the currently popular parameters of f_2/f_1 ratio of 1.2 and intensities of L1−L2=10 dB with L1∼65 dB SPL. However, it is possible that using smaller f_2/f_1 ratios in certain situations could improve the chances of getting a significant response. In 1999, Quinonez (1999) showed maximum DPOAEs to be obtained with an f_2/f_1 of 1.18 using a 5 dB level difference between primaries in pre-term neonates. Although testing at multiple frequency ratios would not be appropriate for screening applications (in the interest of efficiency), we would suggest that varying the f_2/f_1 ratio be considered in follow-up testing of newborns and infants where questionable DPOAE responses are observed under the conventional screening protocol. Similarly, an increase in the f_2/f_1 ratio for the lower frequencies might improve the ability to detect significant DPOAEs in the problematic range below 2 kHz. Such adjustments could potentially reduce false positive findings, for example, in neonatal screening. However, a study focused on outcomes with multiple FRs now is needed to prove the actual improvement in test performance, vis-à-vis sensitivity and specificity and related statistical measures.

Morlet and colleagues (Morlet *et al.*, 1999, 2003) have reported gender and ear differences in TEOAEs testing of pre-term neonates. The ear was not considered in this FR study due to time limitations. A study to determine if ear, as well as gender, differences are important, evident in the developing auditory system's frequency selectivity (as measured by f_2/f_1 functions has not been completed) might yield interesting results to compare with traditional measures of ear dominance established largely in adults.

Finally, to better define and/or test the significance of developmental changes in the auditory function to term birth, a longitudinal design might be preferable, although presenting a formidable challenge for subject recruitment. The work reported here was accomplished via a largely routine newborn test program; parents want to take their babies home as soon as possible. Nevertheless, the potential value of the longitudinal study goes beyond the usual statistical goal (controlling for intersubject variance across age groups). An estimation of the gestational age in infants causes potential overlap in age groups. In pre-term infants, this likely adds variability to the results. As longitudinal designs permit following the same subjects for the observation period, developmental trends might be more salient. As shown by the results reported here, it still would be important to control for gender differences. Testing also should continue through the toddler years. This would allow for an observation of the effects of maturation of the middle ear, to its culmination, while testing auditory function before the environmental effects of noise and toxins might begin to take their toll. Here too, studies emphasizing lower intensity levels might provide other insights on the development of CAMs.

V. CONCLUSION

The results of this study support the results of other researchers showing decreasing optimal ratios for increasing f_2 frequency (Harris *et al.*, 1989; Brown *et al.*, 2000) and suggesting subtle, but continuing, developmental changes in cochlear function through term birth (Abdala, 2003). These changes appear not to be identical between sexes; indeed, collapsing data across gender can potentially blur these trends. This variable certainly should not be ignored in research. Future work in human cochlear development should include measures at lower stimulus levels for pre-term infants, and, ideally, should be of longitudinal design. Results of a longitudinal study would hopefully better describe the pattern of development of cochlear processes than perhaps is revealed by cross-sectional sampling according to estimated gestational ages. Otherwise, the observed changes may simply reflect a sort of flux in the development of uncertain functional significance. While an f_2/f_1 ratio of 1.2 continues to appear broadly applicable, it seems reasonable to investigate whether multiple FR ratios would improve test performance, particularly for purposes of neonatal screening and the goal of reducing costly false positives.

ACKNOWLEDGMENTS

This study was funded by the Research Development Fund of the School of Health and Rehabilitation Sciences of the University of Pittsburgh. I would also like to thank Dr. Catherine Palmer, Dr. Sheila Pratt, Dr. Evelyn Talbot, Dr. Beverly Brezanski, and Dr. Elaine Rubenstein for their assistance in the development and completion of this project.

Abdala, C. (**1996**). ''Distortion product otoacoustic emission ($2f_1-f_2$) amplitude as a function of f_2/f_1 frequency ratio and primary tone level separation in human adults and neonates,'' J. Acoust. Soc. Am. **100**, 3726–3740.

Abdala, C. (**1998**). ''A developmental study of distortion product otoacoustic emission ($2f_1-f_2$) suppression in humans,'' Hear. Res. **121**, 125–38.

Abdala, C. (**2000**). ''Distortion product otoacoustic emission ($2f_1-f_2$) amplitude growth in human adults and neonates,'' J. Acoust. Soc. Am. **107**, 446–566.

Abdala, C. (**2001a**). ''DPOAE suppression tuning: cochlear immaturity in premature neonates or auditory aging in normal-hearing adults?,'' J. Acoust. Soc. Am. **110**, 3155–3162.

Abdala, C. (**2001b**). ''Maturation of the human cochlear amplifier: distortion product otoacoustic emission suppression tuning curves recorded at low and high primary tone levels,'' J. Acoust. Soc. Am. **110**, 1465–1476.

Abdala, C. (**2003**). ''A longitudinal study of distortion product otoacoustic emission ipsilateral suppression and input/output characteristics in human neonates,'' J. Acoust. Soc. Am. **114**, 3239–3250.

Abdala, C., and Chatterjee, M. (**2003**). ''Maturation of cochlear nonlinearity as measured by distortion product otoacoustic emission suppression growth in humans,'' J. Acoust. Soc. Am. **114**, 932–943.

Abdala, C., and Fitzgerald, T. S. (**2003**). ''Ipsilateral distortion product otoacoustic emission ($2f1$-$f2$) suppression in children with sensorineural hearing loss,'' J. Acoust. Soc. Am. **114**, 919–931.

Abdala, C., Ma, E., and Sininger, Y. S. (**1999**). ''Maturation of medial efferent system function in humans,'' J. Acoust. Soc. Am. **105**, 2392–4022.

Abdala, C., and Sininger, Y. S. (**1996**). ''The development of cochlear frequency resolution in the human auditory system,'' Ear Hear. **17**, 374–385.

Abdala, C., Sininger, Y. S., Ekelid, M., and Zeng, F. G. (**1996**). ''Distortion product otoacoustic emissions suppression tuning curves in human adults and neonates,'' J. Acoust. Soc. Am. **98**, 38–53.

Birnholz, J., and Benacerraf, B. (**1983**). ''The development of human fetal hearing,'' Science **222**, 516–518.

Brown, A. M., Sheppard, S. L., and Russell, P. T. (**1994**). ''Acoustic distortion products (ADP) from the ears of term infants and young adults using low stimulus levels,'' Br. J. Audiol. **28**, 273–280.

Brown, A., and Gaskil, S. (**1990**). ''Measurement of acoustic distortion reveals underlying similarities between human and rodent mechanical responses,'' J. Acoust. Soc. Am. **88**, 840–849.

Brown, A., Gaskill, S., Carlyon, R., and Williams, D. (**1993**). ''Acoustic distortion as a measure of frequency selectivity: relation to psychophysical equivalent rectangular bandwidth,'' J. Acoust. Soc. Am. **93**, 3291–3297.

Brown, A., Gaskill, S., and Williams, D. (**1992**). ''Mechanical filtering of sound in the inner ear,'' Proc. R. Soc. London, Ser. A **250**, 29–34.

Brown, A., Sheppard, S., and Russell, P. (**1995**). ''Differences between neonate and adult cochlear mechanical responses,'' Aud. Neurosci. **1**, 169–181.

Brown, D., Kimberly, B., and Eggermont, J. (**1994**). ''Cochlear traveling-wave delays estimated by distortion-product emissions in normal hearing adults and term-born neonates,'' J. Otolaryngol. **23**, 234–238.

Brown, D., Brant, L., and Kimberley, B. (**2000**). ''The effects of maturation and stimulus parameters on the optimal $f(2)/f(1)$ ratio of the $wf(1)$-$f(2)$ distortion product otoacoustic emission in neonates(1),'' Hear. Res. **145**, 17–24.

Chang, K. W., Vohr, B. R., Norton, S. J., and Lekas, M. D. (**1993**). ''External and middle ear status related to evoked otoacoustic emission in neonates,'' Arch. Otolaryngol. Head Neck Surg. **119**, 276–282.

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 amplitudes as a function of $f2/f1$ ratio,'' J. Acoust. Soc. Am. **85**, 220–229.

Kane, J., and Durrant, J. (**1997**). ''Frequency selectivity: comparisons between DPOAE and psychophysical tuning curves in normal-hearing subjects,'' in *abstract of the 20th Midwinter Research Meeting*, A.R.O.

Keefe, D. H., Folsom, R., Gorga, M. P., Vohr, B. R., Bulen, J., and Norton, S. J. (**2000**). ''Identification of neonatal hearing impairment: ear-canal measurements of acoustic admittance and reflectance in neonates,'' Ear Hear. **21**, 443–461.

Keefe, D., Bulen, J., Arehart, K., and Burns, E. (**1993**). ''Ear canal impedance and reflection coefficient in human infants,'' J. Acoust. Soc. Am. **94**, 2617–2638.

Keefe, D., Bulen, J., Campbell, S., and Burns, E. (**1994**). ''Pressure transfer function and absorption cross section from the diffuse field to the human infant ear canal,'' J. Acoust. Soc. Am. **95**, 355–371.

Lasky, R. (**1998a**). ''Distortion product otoacoustic emissions in human newborns and adults. I. Frequency effects,'' J. Acoust. Soc. Am. **103**, 981–991.

Lasky, R. (**1998b**). ''Distortion product otoacoustic emissions in human newborns and adults. II. Level effects,'' J. Acoust. Soc. Am. **103**, 992–1000.

Morlet, T., Durrant, J. D., Lapillonne, A., Putet, G., Collet, L., and Duclaux, R. (**2003**). ''Development of auditory asymmetry in transient evoked otoacousic emissions in pre-term infants,'' J. Am. Acad. Audiol **14**, 339–346.

Morlet, T., Goforth, L., Hood, L. J., Ferber, C., Duclaux, R., and Berlin, C. I. (**1999**). ''Development of human cochlear active mechanism asymmetry: involvement of the medial olivocochlear system? [published erratum appears in Hear. Res. 1999 Nov;137(1-2):179],'' Hear. Res. **134**, 153–622.

Morlet, T., Perrin, E., Durrant, J. D., Lapillonne, A., Ferber, C., Duclaux, R., Putet, G., and Collet, L. (**1996**). ''Development of cochler active mechanisms in humans differs between gender,'' Neurosci. Lett. **220**, 49–52.

Moulin, A. (**2000**). ''Influence of primary frequencies ratio on distortion product otoacoustic emissions amplitude. I. Intersubject variability and consequences on the DPOAE-gram,'' J. Acoust. Soc. Am. **107**, 1460–1700.

Owens, J. J., McCoy, M. J., Lonsbury-Martin, B. L., and Maring, G. K. (**1993**). ''Otoacoustic emissions in children with normal ears, middle ear dysfunction, and ventilating tubes,'' Am. J. Otolaryngol. **14**, 34–40.

Pujol, R., and Lavigne-Rebillard, M. (**1992**). ''Development of neurosensory structures in the human cochlea,'' Acta Oto-Laryngol. **112**, 259–264.

Pujol, R., and Lavigne-Rebillard, M. (**1995**). ''Sensory and neural structures in the developing human cochlea,'' Int. J. Pediatr. Otolaryngol. **32**, S117–S182.

Quinonez, R. E. (**1999**). ''Distortion-product otoacoustic emissions (DPEs) in neonates: frequency ratio (f_2/f_1) and stimulus level differences (L1–L2),'' Acta Oto-Laryngol. **119**, 431–466.

Rubel, E. W. (**1985**). ''Auditory system development,'' in *Measurement of Audition and Vision in the First Year of Postnatal Life: A Methodological Overview*, edited by G. Gottlieb & N. A. Krasnegor. (Ablex, Norwood, NJ), pp. 53–90.

Sato, H., Sando, I., and Takahashi, H. (**1991**). ''Sexual dimorphism and development of the human cochlea,'' Acta Oto-Laryngol. **111**, 1037–1040.

Smurzynski, J. (**1994**). ''Longitudinal measurements of distortion-product and click-evoked otoacoustic emissions of preterm infants: Preliminary results,'' Ear Hear. **15**, 210–223.

Trine, M. B., Hirsch, J. E., and Margolis, R. H. (**1993**). ''The effect of middle ear pressure on transient evoked otoacoustic emissions,'' Ear Hear. **14**, 401–407.

Wada, H., Ohyama, K., Kobayashi, T., Koike, T., and Noguchi, S. (**1995**). ''Effect of middle ear on otoacoustic emissions,'' Audiology **34**, 161–176.

Whitehead, M. L., McCoy, M. J., Lonsbury-Martin, B. L., and Martin, G. K. (**1995**). ''Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. I. Effects of decreasing L2 below L1,'' J. Acoust. Soc. Am. **97**, 2346–2358.

Identification of otoacoustic emissions components by means of adaptive approximations

W. Wiktor Jedrzejczak and Katarzyna J. Blinowska
Laboratory of Medical Physics, Institute of Experimental Physics, Warsaw University, ul. Hoza 69, 00-681 Warszawa, Poland

Wieslaw Konopka
Department of Otolaryngology, Medical University, ul. Zeromskiego 113, 90-549 Lodz, Poland

Antoni Grzanka
Institute of Electronic Systems, Warsaw University of Technology, ul. Nowowiejska 15/19, 00-665 Warszawa, Poland, and Independent Laboratory of Prevention of Environmental Hazards, Medical University, ul. Zwirki i Wigury 61, 02-091 Warszawa, Poland

Piotr J. Durka
Laboratory of Medical Physics, Institute of Experimental Physics, Warsaw University, ul. Hoza 69, 00-681 Warszawa, Poland

(Received 28 October 2003; revised 3 February 2004; accepted 5 February 2004)

Clicks and a set of tone bursts covering the same frequency band were applied as a stimuli evoking otoacoustic emissions (OAE). Recorded otoacoustic emissions were decomposed into the basic waveforms by means of high-resolution adaptive time–frequency approximation method based on the matching pursuit algorithm. The method allows for description of the signal components in terms of frequencies, time occurrences, time spans, and energy. The analysis of OAE's energy density distributions in time–frequency space revealed that click responses can be considered as linear superpositions of responses to tone bursts. The frequency–latency relationship was studied and compared with earlier works. The method made possible the exhaustive description of the resonant modes specific for given subject/ear. They were characterized not only by the close frequencies appearing for different tones, but they usually had similar latencies and time spans. Short-time and long-time resonant modes were identified. The second ones might be connected with spontaneous emissions. The method opens new perspectives in studying the fine structure of the OAE and testing of the theoretical models. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1690077]

PACS numbers: 43.64.Jb [BLM] Pages: 2148–2158

I. INTRODUCTION

Otoacoustic emissions (OAE) were first registered by Kemp (1978) and very quickly became an important tool in the diagnosis of hearing impairment (reviewed by Probst *et al.*, 1991). However, the mechanisms of OAE generation are still a matter of a debate. In particular, it concerns the role of linear and nonlinear effects in shaping the structure of the OAE signal. Spontaneous OAE (SOAE) as well as evoked OAE (EOAE) exhibit periodic variations in amplitude and phase with frequency, which are called ''fine structure.'' Early attempts to describe cochlear OAE fine structure were based on the assumption that they originated from nonlinear reflection (e.g., Kemp, 1997 and the review by Shera and Guinan, 1999). However, there are also models, which consider linear cochlear reflection as a main source of cochlear fine structure. These models involve the presence of low-level inhomogeneities strewn along cochlea (Shera and Zweig, 1993; Zweig and Shera, 1995; Talmadge *et al.*, 2000). In order to resolve the debate, the test of the models based on the precise analysis of the experimental data is needed. In particular, the study of the fine structure of OAE calls for a method which will have the capability of decomposing the signal into basic components of well-defined time–frequency characteristics. This work presents such method, based on adaptive approximations, which has time–frequency resolution superior to the other methods currently applied for OAE studies and at the same time offers parametric description of the signal components.

One of the parameters important for testing the models is the latency of frequency components. The scale-invariance hypothesis (Talmadge *et al.*, 2000) leads to prediction of inverse proportionality between latency and frequency. This hypothesis was conjectured, e.g., by Sisto and Moleti (2002), who pointed out the violation of $1/f$ law. Namely, in some experiments (Tognola *et al.*, 1997) the relation between spectral latency τ and frequency f was fitted by the power law

$$\tau(f) = a f^b, \tag{1}$$

where $a = 10$ ms, $b = -0.4$ (f in kHz). The $\tau(f)$ relation depends strongly on the method of latency estimation.

In the recent studies continuous wavelet transform (WT) (Tognola *et al.*, 1997) or discrete WT (Sisto and Moleti, 2002) was used. However as pointed out in the last reference and also in Moleti and Sisto, 2003, wavelet transform due to its poor frequency resolution, systematically underestimates the slope of the latency–frequency relation.

The method proposed by us—matching pursuit (MP), is free of the limitation of WT, which inversely binds time and frequency bands; moreover, the MP method makes possible the decomposition of the OAE signal into its basic components described by means of well-defined frequencies, latencies, amplitudes, and time spans.

In this paper we shall first describe the method and demonstrate its performance by means of simulations. The comparison with other methods of time–frequency analysis will be made. Then, we will apply MP to the click- and tone-evoked OAE. We shall compare the time–frequency distributions obtained for click-evoked OAE with the superposition of the tone-generated OAE, with the aim of testing for the linearity of the response. Next, we shall construct the frequency–latency curve and fit the function representing this relation. Finally, we will inspect the fine structure of the OAE found by means of the decomposition of the signal into its basic components.

II. METHOD

A. Experimental procedures

Otoacoustic emissions were recorded using the ILO 292 Echoport system designed by Otodynamics. The acquisition window had an onset at 2.5 ms, cosinusoidal rise 2.26 ms, flat top up to 20.5 ms. Dataset consisted of the OAE recordings from 12 young (20–25 years) adult men. All subjects were laryngologically healthy without any otoscopic changes in ears. Impedance audiometry tests for all subjects were performed. Type A tympanograms were recorded with correct reflex from stapedius muscle. In pure-tone audiometry hearing threshold was 10–15 dB. Responses to 260 repetitions of stimuli were averaged with ''nonlinear'' mode of stimulation. Intensity of stimuli was kept on the level of 65–68 dB.

In the study of the responses to tone bursts, stimuli of five frequencies: 1000, 1414, 2000, 2828, and 4000 Hz of half-octave bands were used. These stimuli were constructed to cover the same frequency band (840 to 4757 Hz) as the click stimulus.

B. Data analysis method

The matching pursuit algorithm was introduced by Mallat and Zhang (1993) and first applied to physiological signal processing by Blinowska and Durka (1994). The method relies on adaptive decomposition of the signal into waveforms (also called atoms) from a large and redundant dictionary of functions. A dictionary of basic waveforms can be generated, e.g., by scaling, translating, and unlike in wavelet transform, *modulating* window function $g(t)$

$$g_I(t)=\frac{1}{\sqrt{s}}g\left(\frac{t-u}{s}\right)e^{i\xi t}, \tag{2}$$

where $s>0$=scale, ξ=frequency modulation, u=translation.

In practice for analysis of real-valued signals, the dictionaries are limited to real functions. The index $I=(\xi,s,u)$ describes the set of parameters. The window function $g(t)$ is usually even and its energy in the time domain is mostly concentrated around u with variance proportional to s. In the frequency domain the energy is mostly concentrated around ξ with a spread proportional to $1/s$. The minimum of time–frequency variance is obtained when $g(t)$ is Gaussian.

The dictionaries of windowed Fourier transform and wavelet transform can be derived as subsets of this dictionary, defined by certain restrictions on the choice of parameters. In the case of the windowed Fourier transform, the scale s is constant—equal to the window length—and the parameters ξ and u are uniformly sampled. In the case of WT the frequency modulation is limited by the restriction on the frequency parameter $\xi=\xi_0/s$, $\xi_0=$const.

Finding an optimal approximation of signal by selecting functions from such a large family is a NP-hard problem (computationally intractable). Therefore, a suboptimal iterative procedure is applied. In the first step of the iterative procedure the vector g_{Io} is chosen which gives the largest inner product with the signal $f(t)$

$$f=\langle f,g_{Io}\rangle g_{Io}+R^1f. \tag{3}$$

Then, the residual vector R^1f obtained after approximating f in the direction g_{Io} is decomposed in a similar way. The iterative procedure is repeated on the following obtained residues:

$$R^nf=\langle R^nf,g_{I_n}\rangle g_{I_n}+R^{n+1}f. \tag{4}$$

In this way the signal f is decomposed into a sum of time–frequency waveforms, chosen to match optimally the signal's residues

$$f=\sum_{n=0}^{m}\langle R^nf,g_{I_n}\rangle g_{I_n}+R^{n+1}f. \tag{5}$$

The point at which the iterations should be stopped, or equivalently, the number of waveforms in expansion (5), can be chosen individually for each signal based upon mathematical criteria or set arbitrarily, e.g., as a percentage of energy accounted for. It was proven (Mallat and Zhang, 1993) that the procedure converges to f. Energy of representation is conserved

$$\|f\|^2=\sum_{n=0}^{\infty}|\langle R^nf,g_{I_n}\rangle|^2. \tag{6}$$

The highest time–frequency resolution is obtained for functions g_I from the Gabor family (Mallat, 1999). In this study Gabor functions, sinusoids, and delta functions were used.

One of the problems encountered in the procedure of the adaptive approximations is sampling of the time–frequency space in respect of the fitting atom's positions. In the original algorithm (Mallat and Zhang, 1993) the dyadic sampling was applied; therefore, some positions of atoms were privileged. This effect was observed by us, for averaged distributions of the atom's parameters and also on time–frequency plots. In order to avoid the effects of dictionary structure a new algorithm based on stochastic dictionaries was introduced (Durka *et al.*, 2001a).

We can visualize the results of MP decomposition in the time–frequency plane by adding the Wigner distributions of each of the selected waveforms. Wigner distribution is defined by

$$W[f,h](t,\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty} f\left(t+\frac{\tau}{2}\right)\overline{h\left(t-\frac{\tau}{2}\right)}e^{-i\omega t}d\tau. \quad (7)$$

Calculation of the time–frequency distribution from the whole decomposition produces

$$W[f,f](t,\omega)=\sum_{n=0}^{\infty}|\langle R^n f,g_{I_n}\rangle|^2 W[g_{I_n},g_{I_n}](t,\omega) + \sum_{n=0}^{\infty}\sum_{m=0,m\neq n}^{\infty}\langle R^n f,g_{I_n}\rangle\overline{\langle R^m f,g_{I_m}\rangle} \times W[g_{I_n},g_{I_m}](t,\omega), \quad (8)$$

where the double sum in Eq. (8), containing cross distributions of different waveforms, corresponds to the cross terms generally present in Wigner distribution. One usually tries to remove these terms in order to obtain a clear picture of the energy distribution in the time–frequency plane. Removing these terms from Eq. (8) is straightforward—only the first sum is kept. The energy density in the time–frequency plane of the signal's representation obtained by means of MP is given by the expression $Ef(t,\omega)$

$$Ef(t,\omega)=\sum_{n=0}^{\infty}|\langle R^n f,g_{I_n}\rangle|^2 W[g_{I_n},g_{I_n}](t,\omega). \quad (9)$$

The distribution conserves the signal energy in the time–frequency space

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}Ef(t,\omega)dt\,d\omega=\|f\|^2. \quad (10)$$

This justifies the interpretation of $Ef(t,\omega)$ as the energy density of signal $f(t)$ in the time–frequency plane. All the presentations in this work referred to as time–frequency distributions (also called "t–f maps") are based upon formula (9)—except for the fact that the sum is finite.

The method is very robust in respect to noise. The addition of noise with variance twice bigger than the variance of the signal does not influence critically the time–frequency positions of waveforms corresponding to simulated structures (Blinowska and Durka, 2001). The advantages of the method were demonstrated in the EEG studies, e.g., for extraction of specific structures from the signal (Zygierewicz *et al.*, 1999) and for revealing microstructure of event-related responses (Durka *et al.*, 2001b).

The goodness of fit of the latency–frequency curves was estimated by means of R-square. R-square is defined as the ratio of the sum of squares of the regression and the total sum of squares. R-square can take on any value between 0 and 1, with a value closer to 1 indicating a better fit.

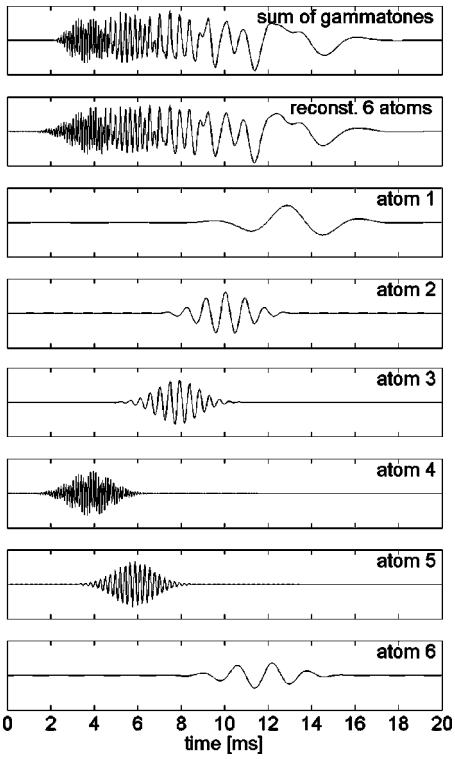


FIG. 1. From the top: the simulated signal—sum of six gammatones of frequencies 280, 550, 1100, 2200, 4400, 8800 Hz, with 2-ms spacing and decay constant of 300 Hz; the reconstruction of the signal from the first six waveforms (atoms), found by the MP procedure; atoms 1–6 shown below.

III. RESULTS

A. Simulations

In order to demonstrate the properties of the MP algorithm and to compare it with other methods used for the evaluation of the OAE signals, we have performed simulations based on the signals resembling OAE. Components of OAE were represented by gammatones, since it is commonly assumed that they reproduce the shape of click-evoked OAE at single resonant frequency. The gammatone is expressed by a waveform, whose envelope grows as t^3 and decays with a constant Γ

$$\gamma(t)=\gamma_0 t^3 e^{-2\pi\Gamma t}\sin(2\pi f t). \quad (11)$$

We have constructed the test signal consisting of six gammatones with frequencies: 280, 550, 1100, 2200, 4400, 8800 Hz, spaced 2 ms in time and decaying with the constant 300 Hz, which gave descending slopes similar to these obtained experimentally. In Fig. 1 the decomposition of this signal by means of the MP algorithm into the six atoms of highest energy is shown, together with the signal reconstructed from these atoms. Correlation coefficient between the reconstructed and original signal is 0.98. The rest of the energy of the original signal is accounted for by the next atoms found in the iterative procedure. They describe the details of the shape of the gammatones; however, their contribution to the total energy of the signal is small. We can see that the positions of the strongest atoms coincide with the maxima of the gamma bursts—basic waveforms well reproduce the components of the simulated signal. Therefore, we can accept cen-

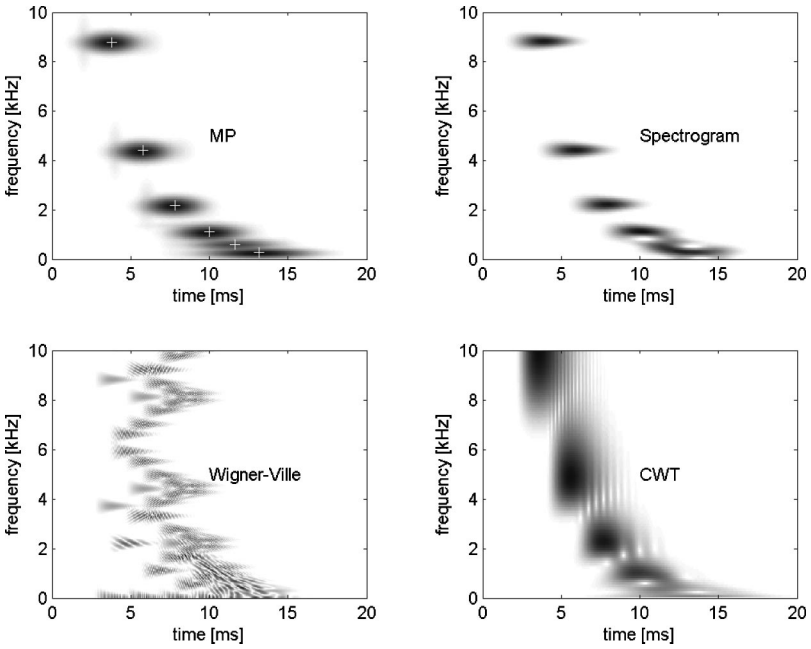


FIG. 2. Time–frequency (t–f) distribution of energy density for simulated signal (shown in Fig. 1) approximated by different methods: MP, spectrogram, Wigner–Ville transform, continuous wavelet transform. White crosses on the first plot indicate t–f centers of the fitted atoms. Energy density proportional to the shades of gray.

ters of the atoms found by MP as the latencies of OAE components connected with "group velocities."

The application of ILO window to the simulated signal resulted in a slight decrease of the amplitude and in the shift toward longer latencies for high-frequency component (0.4 ms for 8800 Hz).

Since we know the parameters of the signal components, we can construct time–frequency distribution of the signal energy according to Eq. (9). In Fig. 2 the time–frequency distributions (t–f distributions), for the simulated signal consisting of six gammatones, found by different methods are shown. The spectrogram gave good results for components of distant frequencies; however, it was unable to separate two lower-frequency components. It is easy to observe the low-frequency resolution of WT for high frequencies and poor time resolution for low frequencies. The Wigner–Ville (W–V) method gives a very poor picture which can be improved by introduction of appropriate corrections decreasing the energy of the cross terms. However, the procedure of improving W–V performance is always to some extent arbitrary; moreover, the W–V method does not provide parametric description of the signal. The same holds for the continuous WT and spectrogram. The MP procedure is characterized by the highest resolution and even more important—it provides the parametric description of the signal components.

The accuracy of the latency and frequency determination by the MP method in relation to the signal simulating OAE was evaluated by the procedure similar to proposed by Sisto and Moleti (2002). The same signal, described above and shown in Fig. 1 (top), was used for this purpose. Additionally, signal with the same frequencies and spacing of the gammatones but decay constant of 600 Hz was used. The relations between frequencies and latencies of gammatones used in simulations and those found by MP are shown in Fig. 3. The spread of the estimation of frequency for 600-Hz decay constant of gammatones is smaller than 30 Hz and for gammatones with the 300-Hz decay constant smaller than 50

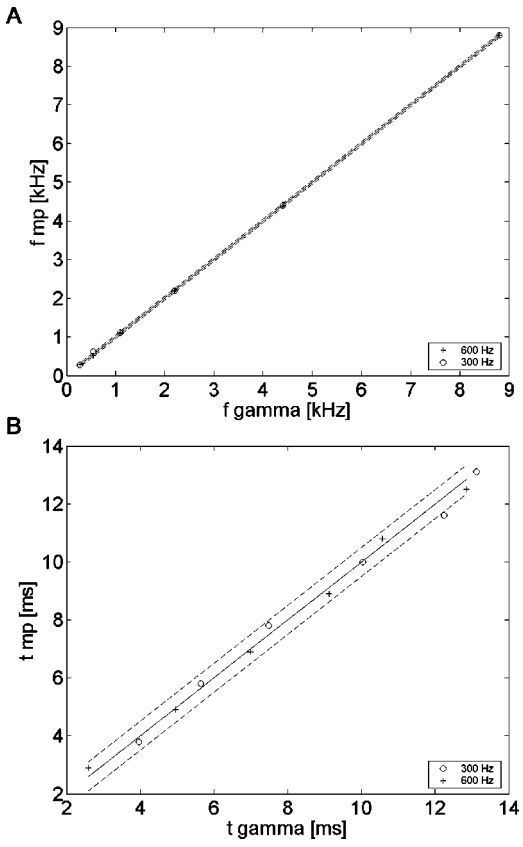


FIG. 3. The test of the accuracy of the frequency determination by the MP method in relation to signals resembling OAE. The same signal as the one shown in Fig. 1 (top) was used (six gammatones of 2-ms spacing, decay constant of 300 Hz—circles) and additionally signal with the same frequencies of gammatones but decay constant of 600 Hz (crosses). (A) the frequency of atoms found by the MP method (vertical axis) is plotted versus the frequency of gammatones. Dashed lines indicate a ± 50-Hz error. (B) time positions of atoms found by the MP procedure (vertical axis) are plotted versus time positions of gammatones.

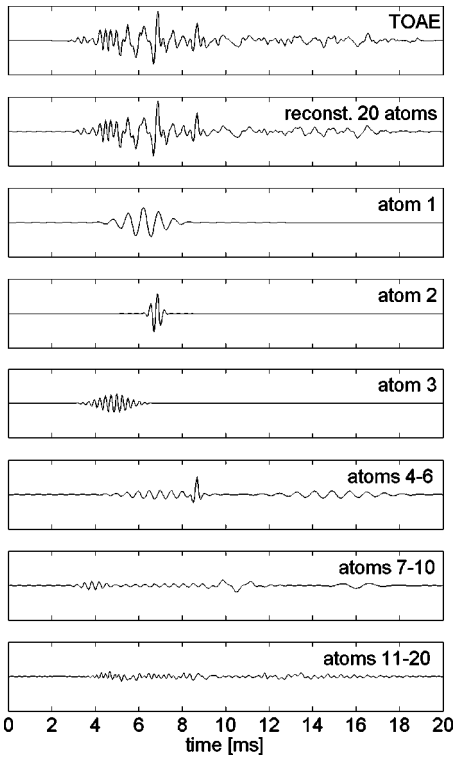


FIG. 4. The MP decomposition of click-evoked OAE (subject "tir"). From top to bottom: original signal, reconstruction from 20 waveforms, waveforms found by the MP algorithm.

Hz. They do not depend on frequency. A similar procedure applied for the determination of the latencies' accuracy indicated the spread of about 0.5 ms.

B. Results for experimental data

Tone- and click-evoked OAE were decomposed by means of the MP algorithm with stochastic dictionaries and the parameters of the components were found. In Fig. 4 an example of the decomposition of click-evoked OAE is shown. The basic features of the OAE are reproduced by the first 15 atoms, which account for 95% of the energy of the signal. When the components of the signal are known, it is straightforward to construct the time–frequency distribution of the energy density (Fig. 5).

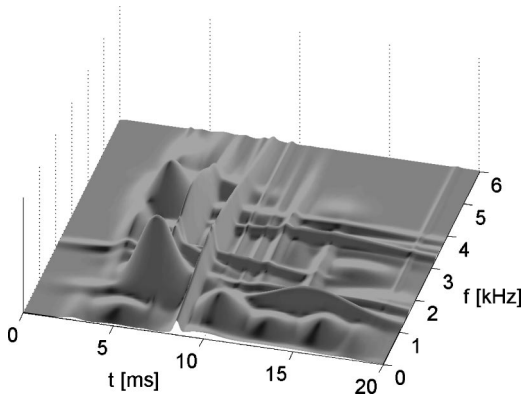


FIG. 5. Time–frequency distribution of energy density (sqrt scale, proportional to height) of click-evoked OAE (the same signal as in Fig. 4) obtained by means of MP. Subject "tir."

The time–frequency (t–f) representations of the tone-evoked OAE are presented in Fig. 6. Usually the atom of the highest energy was closest to the frequency of the stimulation. In order to compare tone- and click-evoked OAE we have made a t–f map constructed from the sum of the energy distributions of responses to the different tones. In Fig. 7 such t–f map is shown together with the t–f distribution obtained for click. The t–f maps of the corresponding stimuli are also enclosed. It is easy to see that the sum of the energy distributions of the tone-evoked OAE corresponds rather well to energy distribution of the click OAE. In order to make the comparison easier, in the next figure (Fig. 8) the centers of the atoms for click- and tone-evoked OAE are shown together. We can observe that the centers of tone-evoked OAE components in t–f space tend to be shifted toward longer latencies. This might have been expected from the fact that the stimuli in the case of bursts were applied with some delays (Fig. 7).

In the previous considerations (Figs. 7 and 8) the examples of t–f maps of OAE energy density for one subject are shown. The inspection of the t–f distributions for all subjects revealed the same effects. Summary of the results for all the subjects is shown in Fig. 9. We can see that click- and tone-evoked responses span the similar part of the time–frequency space. The atoms corresponding to tones are slightly shifted toward longer latencies; however, this is due to the t–f structure of the stimuli (compare Fig. 7). This result indicates that to a large extent click stimulation can be considered as a superposition of tone responses, which points toward involvement of linear mechanism, at applied level of stimulation.

The MP method provides directly the latencies and frequencies of the OAE components; therefore, construction of the latency–frequency dependence is straightforward. In the case of tone-evoked OAE the positions of atoms did not always correspond exactly to the frequencies of stimulation; therefore, the highest energy atoms occurring within ± 0.25 octave band, in respect to the stimulation frequency, were used for the construction of the frequency–latency curve. For the click-evoked OAE five highest energy atoms were selected.

We have fitted different kinds of power-law functions to represent at best the frequency–latency relationship for the click- and tone-evoked OAE. The best fit for tone-evoked OAE was obtained for the function

$$\tau(f) = a f^b, \tag{12}$$

with parameters $a = 11.5$ ms, $b = -0.62$. The goodness of fit measured by the R-square was 0.74. The corresponding curve is shown in Fig. 10. The latencies of the OAE are influenced by the characteristic of stimulus and the ILO window. The characteristics of tone stimuli shift the latencies of tone-evoked OAE toward larger values. ILO window can increase the latencies of highest frequency components as well. The exact numerical estimation of both effects is difficult and in the case of apparatus window based only on simulations. We decided against introduction of corrections, since they would be a source of additional errors. However, the above-mentioned effects influence the results of curve

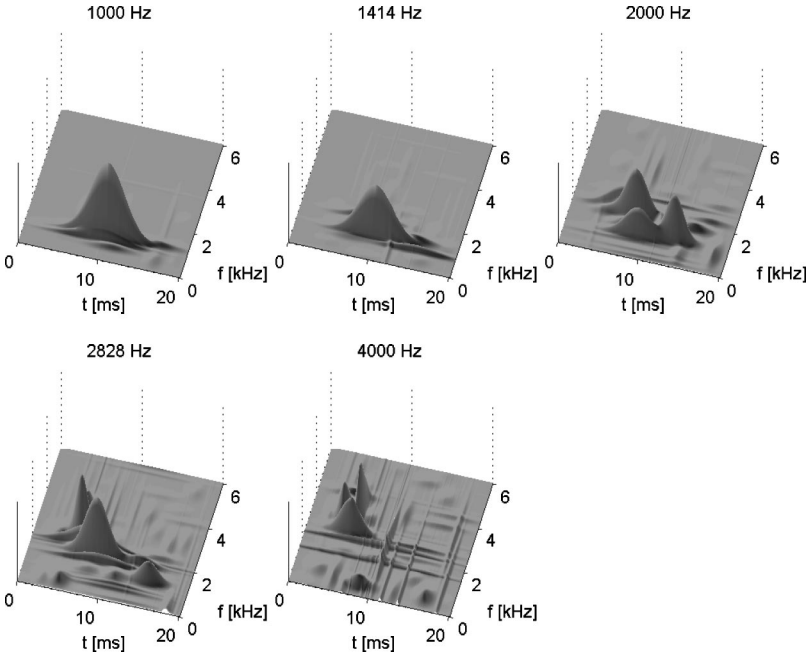


FIG. 6. Time–frequency distributions of energy density of the responses for tone-burst stimuli (from 1000 to 4000 Hz) obtained from the MP decomposition. Frequency of stimulation given above the maps. On each plot energy scaled to the maximal value. Subject "skr."

fitting and obtained parameters have to be treated with reserve.

The fit was performed also for click-evoked OAE (Fig. 11). In this case the parameters of the function given by Eq. (12) were: $a = 10.82$ ms, $b = -0.565$ (R-square: 0.44). In Fig. 11 also the fits obtained for functions proposed by Sisto and Moleti (2002) and the $1/f$ relation suggested by Talmadge (2000) are shown.[1)]

Superior time–frequency resolution of the presented method made possible the study of the relationship between the stimulus and the pattern of the time–frequency cochlear response. The resonant modes in OAE were recognized earlier on the basis of spectral information (e.g., Probst *et al.*, 1991). The MP method provides the information not only on their frequency, but also on their latencies and time spans. The method is robust to noise (Blinowska and Durka, 2001), and even weak resonant modes can be unequivocally identified. An example of the resonant modes found by the algorithm is illustrated in Fig. 12. For stimulation frequency of 1414 Hz the second largest atom has the frequency 1790 Hz, which is close to the strongest component (1815 Hz) for 2000-Hz excitation. The component of almost the same frequency appears as well, as a second atom (1855 Hz) for stimulation at 2828 Hz. The component of frequency around 2450 Hz appears for stimulations at 2000, 2828, and 4000 Hz as a third, first, and fifth atom, correspondingly (the atoms are numbered according to their energy, starting with the strongest one).

The atoms of close frequencies appearing for different tones have similar latencies, which confirms the fact that

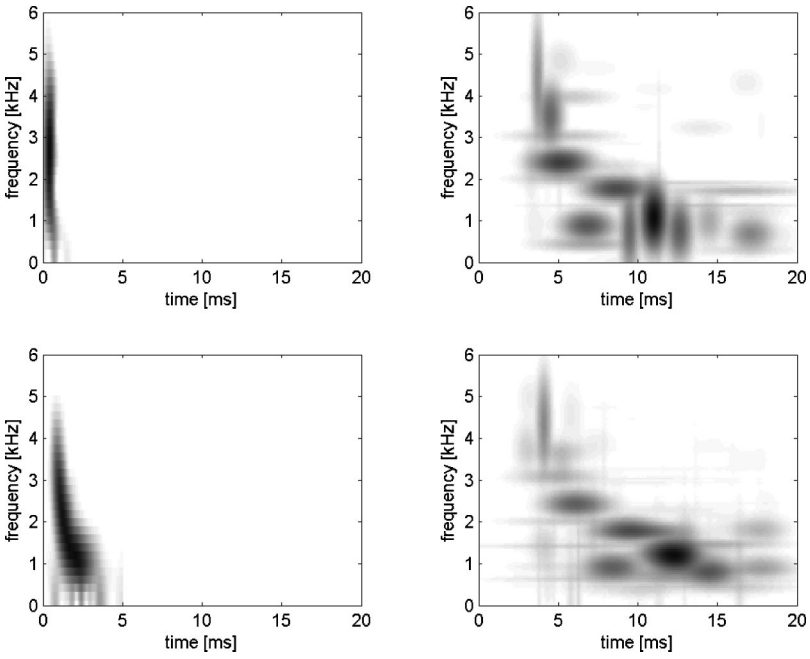


FIG. 7. Time–frequency distributions of energy. On the left: t–f maps of stimuli: top—click, bottom—sum of the maps of bursts. On the right: t–f maps of OAE response to click (top) and sum of the maps of responses to tones (bottom). The shades of gray correspond to the energy density of the signal. On each plot energy scaled to the maximal value. Subject "tfr."

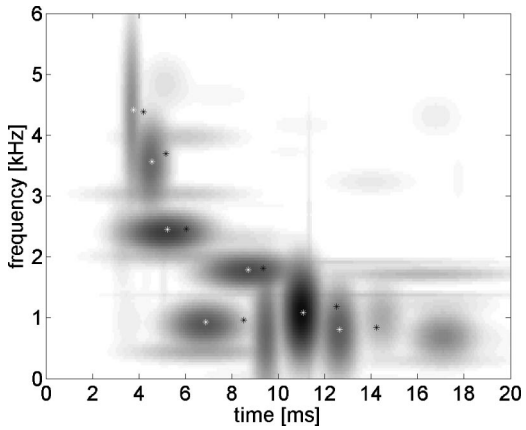


FIG. 8. Time–frequency distribution of energy obtained by means of the MP decomposition of click-evoked OAE. White stars indicate t–f centers of the main atoms of click-evoked OAE. Black stars mark the positions of the strongest atoms of responses to tone bursts. Subject ''tfr.''

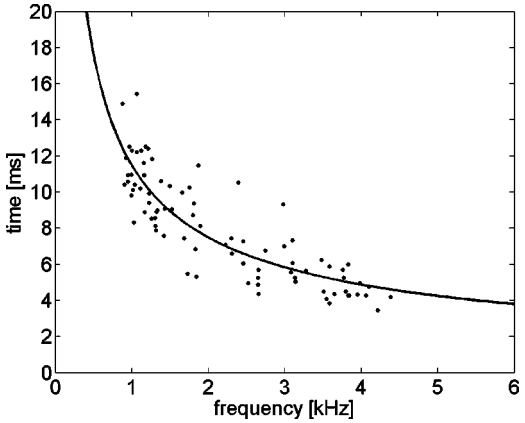


FIG. 10. Centers of waveforms in time–frequency space fitted by the MP procedure to tone-evoked OAE for all subjects. Since the positions of atoms did not correspond exactly to the frequencies of stimulation, the highest energy atoms occurring within ± 0.25 octave band in respect to the stimulation frequency were selected. The curve $t = af^b$ fitted to the points with parameters $a = 11.5$ ms, $b = -0.62$ (R-square: 0.74).

they are indeed the OAE components, not the artifacts of our method of decomposition. The described effect appeared for all subjects. Each subject had his own privileged resonance frequencies and these frequencies were different for the right and for the left ear. Moreover, for the given ear the atoms corresponding to the same resonant modes usually have similar time spans, but these time spans varied for different resonant modes (the span of the atom is defined as a half-width of the Gabor function envelope).

In the identification of resonant modes not only similarity of frequencies, but also latencies and time spans were considered. As a guideline the accuracy of the frequency and latency determination obtained by means of simulations (Fig. 3) was used. Only the atoms of highest energy were taken into account—mostly atoms up to the fourth iteration (related to the relative energy), only a few atoms of lower energies (up to 7) were taken into account.

The varied individual features of resonant modes made the construction of the statistic measures difficult. In Table I the resonance frequencies for eight ears (the full table would be excessively long) are given together with the annotation concerning the frequency of stimulus, the ''number of the atom,'' which reflects the relative energy of the component, and the time span of the atom.

With the aim of summarizing the results we have constructed a histogram (Fig. 13) showing the number of atoms, which we consider to be resonant modes, as a function of the frequency and latency differences. In Fig. 13 also some cases, which were outside one of the limits of accuracy suggested by the simulations (namely 100 Hz and 2 ms), were included, when the other parameter showed very small discrepancy.

Among the cases shown in Fig. 13 ten subjects have more than one resonance frequency in one of the ears. For one of the subjects the resonant modes occurred only in one ear. The maximal number of resonant modes found in one ear was four. The average number of resonant modes per ear ($\pm$ its standard deviation) was 2 ± 0.9. Mean frequency difference between modes was 41 ± 35 Hz, mean difference between latencies: 1.1 ± 1.1 ms. If we apply the more conservative limits for selecting the resonance modes (Δf

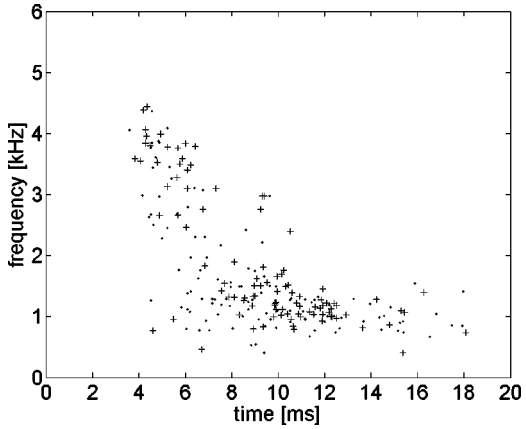


FIG. 9. Positions of centers of atoms in time–frequency space for all subjects (24 ears) in both types of stimulation. Dots indicate t–f centers of the five strongest atoms fitted to the click-evoked OAE of each of the subjects. Crosses mark the strongest atom of MP decomposition of response to tone burst (five tone bursts of different frequencies for each person).

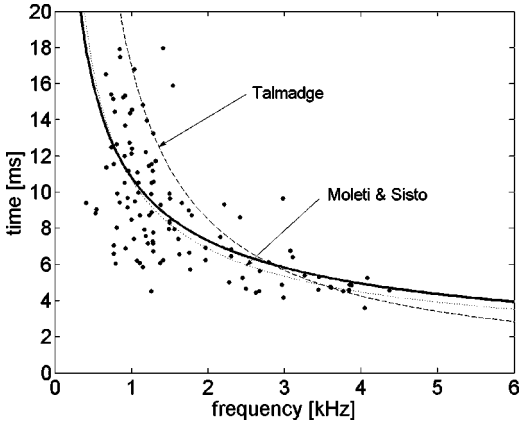


FIG. 11. Centers of waveforms in t–f space fitted by the MP algorithm to click-evoked OAE responses for all subjects. Five highest energy atoms of MP decomposition were selected. The solid curve is a power $t = af^b$ fit to the data with parameters $a = 10.82$ ms, $b = -0.56$ (R-square: 0.44). Dotted curves correspond to the functions proposed by Sisto and Moleti (2002)—$t = 10.2\, f^{0.68} + 0.5$ and the $17/f$ relation suggested by Talmadge (2000).

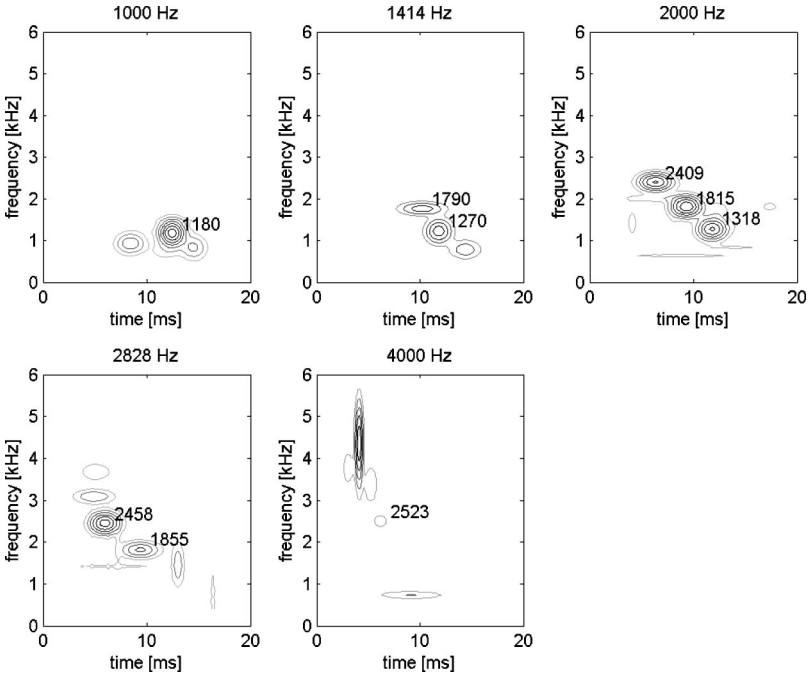


FIG. 12. Contour distributions of energy in time–frequency space for OAE signal evoked by tone-burst stimuli (from 1000 to 4000 Hz). Frequency of the stimulation is given above the maps. Frequencies of atoms with similar t–f parameters, which appear for different stimulus frequencies, are written next to them (in Hz). Subject ''tfr.''

<100 Hz and $\Delta t<3$ ms) the average number of resonance modes will be 1.9 ± 1.0.

In Fig. 14 the histogram of the time spans of the resonant modes is shown. It has a bimodal character. The long-time resonant modes can be connected with spontaneous emissions (Sisto *et al.*, 2001).

IV. DISCUSSION AND CONCLUSIONS

The MP method based on adaptive approximations proved to be a useful tool in the study of OAE. It allowed us not only to identify OAE intrinsic components, but it provided their detailed description in terms of their latencies and time spans. The time–frequency resolution of the MP method is superior to windowed Fourier transform or wavelet transform. Contrary to WT, the MP method does not assume any arbitrary frequency bands. It does not require, as is the case for Wigner–Ville or Choi–Williams transforms, introduction of corrections connected with cross terms, which is always to some extent a subjective procedure. The property which distinguishes MP from other methods is the description of the components of a signal by means of parameters of the clear meaning, namely: their latencies, frequencies, time spans, and energy (or amplitude). Usually most of the energy of the signal is described by a few components only.

In the past the problem of the identification of the latencies of the frequency components of OAE signal was approached in different ways: Fourier transform and discrete or continuous WT were used. The method based on combination of spectral analysis and WT was also proposed (Sisto and Moleti, 2002). The main limitation of the WT is the octave-band resolution, which influences the accuracy especially for high frequencies. In order to surmount this diffi-

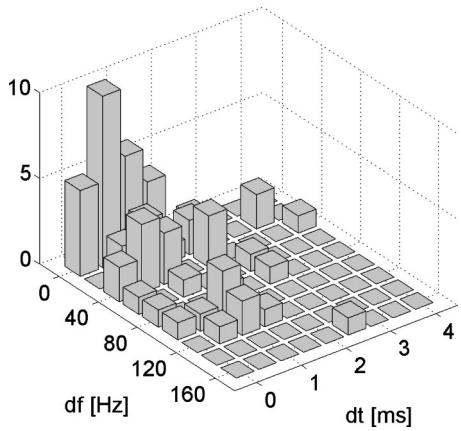


FIG. 13. Histogram of differences between frequencies and latencies of the signal components, which might be considered as resonance modes. The cases of atoms of close frequencies and latencies occurring for at least two different tones were considered. At the axis ''df'' differences between frequencies in Hz, at the axis ''dt,'' differences in latencies in ms.

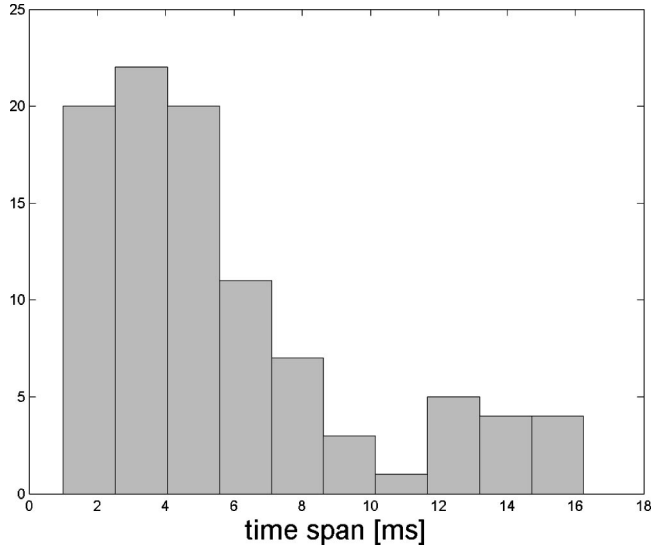


FIG. 14. Histogram of the time spans of the atoms which can be considered as resonant modes. Time span is defined as the half-width of the Gabor function. The same cases as the ones represented in Fig. 13 are shown.

TABLE I. Resonance modes found for four subjects. The last letter in subject acronym indicates from which ear the emissions were measured: l—left, r—right. Different resonance frequencies are separated by horizontal lines. Second and third columns give t–f position of the component. Last column indicates the number of iteration, in which given atom was fitted in the MP decomposition, which corresponds to the relative energy of the component.

Subject	Resonant freq. [Hz]	Latency [ms]	Time span [ms]	Stimulus freq. [Hz]	Nr of atom
tfl	1180	11	3.56	1000	1
	1220	10.76	2.92	1414	1
	1782	9.08	3.72	1414	2
	1863	7.92	4.52	2000	1
	1831	9.04	2.8	2828	2
	2780	6.44	2.32	2828	1
	2791	6.4	2.36	4000	2
tfr	1180	12.52	2.16	1000	1
	1270	11.84	2.28	1414	1
	1318	11.84	2.52	2000	2
	1790	10.28	4.12	1414	2
	1814	9.36	2.64	2000	1
	1855	9.44	3.28	2828	2
	2458	6.04	2.56	2828	1
	2523	6.24	2.24	4000	5
mpgl	1383	10.6	2.6	1414	1
	1383	10.6	1.76	2000	1
mpgr	1310	10.48	5.64	1000	3
	1334	9	1.16	1414	1
skl	1041	10.44	4.52	1000	1
	1041	10.88	3.32	1414	1
	1521	9.04	1.72	1414	2
	1570	8.64	1.24	2000	2
skr	1114	10.2	6.28	1000	1
	1114	9.92	6.16	1414	1
	1074	9.1	9.08	2000	3
	1245	16.24	16.24	1000	3
	1277	14.68	14.68	1414	3
	1261	13.4	13.4	2000	2
	1896	8.12	8.12	2000	1
	1912	8.2	8.2	2828	2
	2661	5.68	5.68	2828	1
	2718	5.36	5.36	4000	2
wal	1261	14.64	3.52	1414	2
	1277	14.24	4.2	2000	1
	1660	9.96	2.96	1414	1
	1676	10	3.44	2000	2
	1692	10.16	3.44	2828	1
	2042	7.2	4.08	2000	6
	2075	7.12	4.4	2828	5
war	1505	8.96	4.84	1414	1
	1505	9.24	2.72	2000	1

culty the method based on the identification of consecutive cochlear echoes was proposed (Sisto and Moleti, 2002). The method relied on visual comparison between wavelet data and transiently evoked OAE spectra, followed by identification of the wavelet contribution due to a given individual spectral line. However, as the authors admit themselves, the method can be applied only to the subset of data for which an unambiguous line identification is possible, and this se-

lection can bias the results. The MP method allows for straightforward and objective identification of the latencies of the frequency components.

Comparison of t–f energy distributions for tone- and click stimuli (for the level of stimuli applied in this study) revealed that the click-evoked responses to a large extent correspond to the superposition of tone-burst-evoked responses, which indicates the similar mechanisms of the generation in both regimes of stimulation and a minor influence of the cochlear nonlinearities. However, latency–frequency function predicted by linear theory (Talmadge *et al.*, 2000) does not give the best fit to the data. It may be connected with the fact that theoretical predictions refer to cochlear latency only, while the OAE latency includes also a noncochlear contribution from the middle ear. The effects connected with the apparatus window may also influence the character of the function. The influence of the window on systematic latency error was studied by Moleti and Sisto, 2003. We have also tried to take this effect into account; however, it may be evaluated only approximately by means of simulations. It seems that some more theoretical and experimental work is needed to explain the character of frequency–latency dependence and to elucidate the OAE generation.

The presented method can be especially suitable in this respect. In particular, it can help in elucidating the presence of the effects connected with the fine structure of OAE. The observation that for different frequencies of stimulation the same preferred response frequencies appeared in OAE spectrum was made already by Elberling *et al.* (1985). However, the Fourier method available at that time did not allow one to obtain time–frequency characteristics of these preferred frequencies. Already in the early eighties the models were proposed with the aim of explaining the privileged frequencies in the OAE signals (e.g., Manley, 1983; Sutton and Wilson, 1983; Neely and Norton 1987), which assumed the presence of irregularities along the cochlear partition.

More recently the OAE fine structure was interpreted in terms of linear cochlear reflection (Zweig and Shera, 1995; Shera and Guinan, 1999; and Talmadge *et al.*, 1998). In particular, Zweig and Shera (1995) postulated the role of low-level inhomogeneities strewn along cochlea, and Talmadge *et al.* (2000) considered the sources of cochlear reflections connected with roughness and nonlinearity. The proposed method is an ideal tool to study the fine structure of the OAE and to test the theoretical predictions. It brings additional information about the signal components not available by the other methods.

In this paper we have demonstrated the capabilities of the MP method, which opens the way to the large repertoire of further studies. The representation of the signal as a superposition of the well-defined components will help in revealing the regularities in the fine structure of the OAE. This in turn can lead to the rules governing the generation of the particular frequencies. This information might be crucial in testing the theoretical models of the OAE generation.

The method also will be very helpful in establishing the connections between spontaneous emissions and the evoked OAE. Their time–frequency patterns can be described accurately in terms of well-defined components and common features can be extracted.

The advantages of the MP method demonstrated in this paper allow us to conjecture that the introduction of the adaptive approximations to the OAE signal may open quite new possibilities in the understanding of the mechanisms of the OAE generation. Moreover, MP can find application in analysis of different acoustic signals, especially when fine time–frequency structure is of interest.

ACKNOWLEDGMENT

This work was partly supported by a grant of the Polish Committee for Scientific Research (KBN) to the Institute of the Experimental Physics of Warsaw University.

[1]In Figs. 10 and 11, the time–frequency axes are reversed in comparison to the earlier pictures in order to make the pictures compatible with the other works, although authors are convinced that it is more natural to have a horizontal time axis.

Blinowska, K. J., and Durka, P. J. (**1994**). ''The Application of Wavelet Transform and Matching Pursuit to the Time-Varying EEG Signals,'' in *Intelligent Engineering Systems through Artificial Neural Networks*, edited by C. H. Dagli and B. R. Fernandez (ASME, New York) Vol. 4, pp. 535–540.

Blinowska, K. J., and Durka, P. J. (**2001**). ''Unbiased high resolution method of EEG analysis in time–frequency space,'' Acta Neurobiol. Exp. (Warsz) **61**, 157–174.

Durka, P. J., Ircha, D., and Blinowska, K. J. (**2001a**). ''Stochastic time-frequency dictionaries for Matching Pursuit,'' IEEE Trans. Signal Process. **49**, No. 3, 507–510.

Durka, P. J., Ircha, D., Neuper, Ch., and Pfurtscheller, G. (**2001b**). ''Time–frequency microstructure of event-related EEG desynchronization (ERD) and synchronization (ERS),'' Med. Biol. Eng. Comput. **39**, No. 3, 315–321.

Elberling, C., Parbo, N. J., Johnsen, N. J., and Bagi, P. (**1985**). ''Evoked acoustic emissions: Clinical application,'' Acta Oto-Laryngol., Suppl. **421**, 77–85.

Kemp, D. T. (**1978**). ''Stimulated acoustic emissions from within the human auditory system,'' J. Acoust. Soc. Am. **64**, 1386–1391.

Kemp, D. T. (**1997**).

Mallat, S. G., (**1999**). *Wavelet Tour of Signal Processing* (Academic, New York).

Mallat, S. G., and Zhang, Z. (**1993**). ''Matching pursuit with time-frequency dictionaries,'' IEEE Trans. Signal Process. **41**, 3397–3415.

Manley, G. A. (**1983**). ''Frequency spacing of acoustic emissions: A possible explanation,'' in *Mechanisms of Hearing*, edited by W. R. Webster and L. M. Aitkin (Monash University Press, Clayton, Australia), pp. 36–39.

Moleti, A., and Sisto, R. (**2003**). ''Objective estimates of cochlear tuning by otoacoustic emission analysis,'' J. Acoust. Soc. Am. **113**, 423–429.

Neely, S. T., and Norton, S. J. (**1987**). ''Tone-burst-evoked otoacoustic emissions from normal-hearing subjects,'' J. Acoust. Soc. Am. **81**, 1860–1872.

Probst, R., Lonsbury-Martin, B. L., and Martin, G. K. (**1991**). ''A review of otoacoustic emissions,'' J. Acoust. Soc. Am. **89**, 2027–2067.

Shera, C. A., and Guinan, Jr., J. J. (**1999**). ''Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy of mammalian OAEs,'' J. Acoust. Soc. Am. **105**, 782–798.

Shera, C. A., and Zweig, G. (**1993**).

Sisto, R., Moleti, A., and Lucertini, M. (**2001**). ''Spontaneous otoacoustic emissions and relaxations dynamics of long decay time OAEs in audiometrically normal and impaired ears,'' J. Acoust. Soc. Am. **111**, 297–308.

Sisto, R., and Moleti, A. (**2002**). ''On the frequency dependence of the otoacoustic emission latency in hypoacoustic and normal ears,'' J. Acoust. Soc. Am. **111**, 297–308.

Sutton, G. J., and Wilson, J. P. (**1983**). ''Modeling cochlear echoes: The influence of irregularities in frequency mapping on summed cochlear activity,'' in Mechanics of Hearing, edited by E. de Boer and M. A. Viergever (Delft University Press, Delft, The Netherlands), pp. 83–90.

Talmadge, C. L., Tubis, A., Long, G. R., and Piskorski, P. (**1998**). ''Modeling otoacoustic emission and hearing threshold fine structures,'' J. Acoust. Soc. Am. **104**, 1517–1543.

Talmadge, C. L., Tubis, A., Long, G. R., and Tong, C. (**2000**). ''Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure,'' J. Acoust. Soc. Am. **108**, 2911–2932.

Tognola, G., Grandori, F., and Ravazzani, P. (**1997**). ''Time-frequency distributions of click-evoked otoacoustic emissions,'' Hear. Res. **106**, 112–122.

Zweig, G., and Shera, C. A. (**1995**). ''The origin of periodicity in the spectrum of evoked otoacoustic emissions,'' J. Acoust. Soc. Am. **98**, 2018–2047.

Zygierewicz, J., Blinowska, K. J., Durka, P. J., Szelenberger, W., Niemcewicz, Sz., and Androsiuk, W. (**1999**). ''High resolution study of sleep spindles,'' Clin. Neurophysiol. **110**(12), 2136–2147.

Cochlear hysteresis: Observation with low-frequency modulated distortion product otoacoustic emissions

Lin Bian,[a] Erin E. Linhardt, and Mark E. Chertoff
Department of Hearing and Speech, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, Kansas 66160

(Received 21 November 2003; revised 26 January 2004; accepted 30 January 2004)

Low-frequency modulation of distortion product otoacoustic emissions (DPOAEs) can be used to estimate a nonlinear transducer function (f_{Tr}) of the cochlea. From gerbils, DPOAEs were measured while presenting a high-level bias tone. Within one period of the bias tone, the magnitudes of the cubic difference tone (CDT, $2f_1-f_2$) demonstrated two similar modulation patterns (MPs) each resembled the absolute value of the third derivative of the f_{Tr}. The center peaks of the MPs occurred at positive sound pressures for rising in bias pressure or loading of the cochlear transducer, and more negative pressures while decreasing bias amplitude or unloading. The corresponding f_{Tr} revealed a sigmoid-shaped hysteresis loop with counterclockwise traversal. Physiologic indices that characterized the double MP varied with primary level. A Boltzmann-function-based model with negative damping as a feedback component was proposed. The model was able to replicate the experimental results. Model parameters that fit to the CDT data indicated higher transducer gain and more prominent feedback role at lower primary levels. Both physiologic indices and model parameters suggest that the cochlear transducer dynamically changes its gain with input signal level and the nonlinear mechanism is a time-dependent feedback process. [DOI: 10.1121/1.1690081]

PACS numbers: 43.64.Jb, 43.64.Kc, 43.64.Bt [BLM] Pages: 2159–2172

I. INTRODUCTION

It has long been speculated (Gold, 1948; Davis, 1983) that an ''active process'' within the inner ear is responsible for the high sensitivity and frequency selectivity in hearing. The finding that the inner ear can emit acoustic energy (Kemp, 1978), known as otoacoustic emissions (OAEs), could indicate ''active'' force productions within the cochlea. A later observation on isolated mammalian cochlear outer hair cells (OHCs) identified that a voltage-dependent motility may be the source for the force production (Brownell *et al.*, 1985). The dominant efferent-nerve innervation (Spoendlin, 1985) and motility suggest that the role of the OHC is to provide a feedback mechanism to enhance the vibration of basilar membrane (BM) and amplify the mechanical stimulation to inner hair cells (IHCs), the primary afferent sensory cells. Coupled with the BM and tectorial membrane, the OHCs form the basic structural and functional units, the bidirectional cochlear transducers. The forward mechanoelectrical transduction (MET) is the conversion of BM motion to voltage change by the gating of ion channels in the hair bundles (Hudspeth, 1989). In mammals, the reverse electromechanical transduction is mediated by the voltage-dependent conformational change of motor proteins, e.g., prestin (Liberman *et al.*, 2002), residing in the basolateral membrane of the OHC. Both transduction processes follow a common saturating nonlinearity [Fig. 1(a)] of Boltzmann-type (Hudspeth, 1989; Evans *et al.*, 1991). Estimating the operating characteristics of the cochlear transducer (transducer function, f_{Tr}) provides a means of evaluating cochlear functional status.

When stimulated by two tones ($f_1, f_2, f_1<f_2$), the cochlea can produce a series of distortion products (DPs) that are detectable in the ear canal (Kemp, 1979). These DPOAEs have been shown to be consistent with a single generating mechanism (Lukashkin *et al.*, 2002; Mills, 2002) that is the nonlinearity in cochlear transduction (Lukashkin and Russell, 1999; Fahey *et al.*, 2000). Recently, Bian *et al.* (2002) found that at low signal levels the magnitude of the cubic difference tone (CDT, $2f_1-f_2$) is proportional to the absolute value of the third derivative of a sigmoid-shaped nonlinear function, denoted as $|f_{Tr}^{(3)}|$. A subsequent experiment in gerbils verified such a relation between CDT and the f_{Tr}. This was done by introducing a low-frequency bias tone to shift a two-tone signal through the operating range of the f_{Tr} in multiple steps, and measuring CDT amplitudes accordingly. The CDT magnitude as a function of the bias level demonstrated a shape similar to the $|f_{Tr}^{(3)}|$ [Fig. 1(d)]. Curve fitting allowed the derivation of the f_{Tr} in the form of a second-order Boltzmann function (f_B):

$$y(x)=\frac{A}{1+e^{b(x+c)}\cdot[1+e^{d(x+e)}]}, \tag{1}$$

where x is bias level in Pa; y is mechanical response in Pa; and A, b, c, d, and e are constants.

The cochlea is a system of distributed nonlinear transducers each tuned to a specific frequency. Since the two tones are high frequency signals, the transducer evaluated with this low-frequency biasing method is located at a relatively basal region on the cochlear partition tuned to the frequency between f_1 and f_2, or closer to f_2 (Martin *et al.*, 1987). For these basal locations, the BM vibration in re-

[a]Electronic mail: lbian@kumc.edu

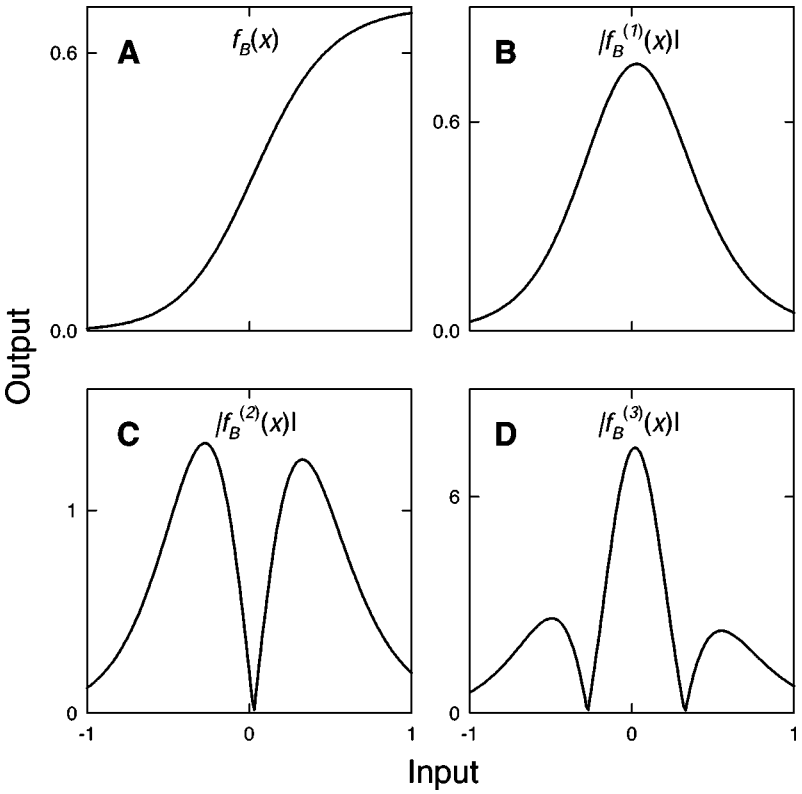


FIG. 1. A second-order Boltzmann function (f_B, panel a) and the absolute values of its first, second and third derivatives (panels b–d).

sponse to the low-frequency bias tone is linear and in-phase with the stimulus (for review, see Robles and Ruggero, 2001), i.e., the sound pressure level of the bias tone [x in Eq. (1)] is equivalent to the BM displacement. Moreover, the observations, that low-frequency modulation of the BM response at characteristic frequency (CF) is an effect of the BM displacement due to the low-frequency tone (Cooper, 1996; Geisler and Nuttall, 1997), also support this equivalence about the x-axis. Assuming a constant cross-sectional area of the ear canal, the output sound pressure [y in Eq. (1)] of the f_{Tr} is proportional to the force generated within the transducer in response to the stimuli. The shape of the f_{Tr} reflects the change in force production as a function of the BM displacement due to the cyclic loading of the cochlear transducer by the bias tone. Therefore, the f_{Tr} is essentially a force-displacement or ''stress–strain'' relationship of the cochlear transducer measured noninvasively from the ear canal. Analysis of the force-displacement relation can reveal the nature of the dynamical behaviors of mechanical systems (Bendat *et al.*, 1992). To study the dynamics of the cochlear transducer on a cycle-by-cycle basis, a procedure was developed to evaluate the temporal and mechanical characteristics of the low-frequency modulated DPOAEs.

II. METHODS

A. Experimental procedures

The experimental procedures were approved by the Institutional Animal Care and Use Committee at the University of Kansas Medical Center. Ten Mongolian gerbils (*Meriones unguiculatus*), weighing between 52 and 75 g, were used as experimental subjects. After anesthetized with Ketamine/Xylazine (100 and 2 mg/kg, i.p.), and local injection of 1% lidocaine, the right pinna was removed and a speculum was sealed onto the bony ear canal. A calibrated probe microphone (Etymotic Research, ER-10B) was inserted into the speculum. The body temperature was maintained at 37 °C with a heating pad (Harvard). To ensure stability and reliability of the experiment, the CDT magnitude was monitored on a spectral analyzer (Hewlett-Packard, 3561A) with the primary levels set at 80 dB SPL ($L_1=L_2$) without bias tone. This CDT magnitude was maintained in the normal range ($\geqslant$40 dB SPL) prior to each recording session throughout the experiment. Then, the ear-canal acoustics were recorded while presenting a two-tone signal and a bias tone. There was a recording session at each of the five different primary levels for two pairs of primary tones. Next, the whole procedure was repeated with an interval of about 90 min. Each of the two repetitions is termed a trial.

B. Signal processing and data analysis

A 25-Hz bias tone was presented at 120 dB pSPL (20 Pa peak amplitude) to modulate the DPOAEs, since it could provide the range of BM displacement necessary for deriving the cochlear f_{Tr}. This method was more efficient than the previous procedures (Bian *et al.*, 2002), because data only needed to be collected for one bias tone, instead of 21 bias tones with various amplitudes. By selecting segments of the acoustic signal at peaks and troughs of the bias tone, the previous method accessed the CDT magnitudes at fixed phases of the bias tone and eliminated any temporal or dynamical effect. Portions of the results (seven animals) were from reanalyses of the previous data. Two pairs of primary frequencies were 3968/5120 Hz ($f_2/f_1=1.29$) and 5888/7552 Hz ($f_2/f_1=1.28$) with levels ranging from 50 to 70 dB

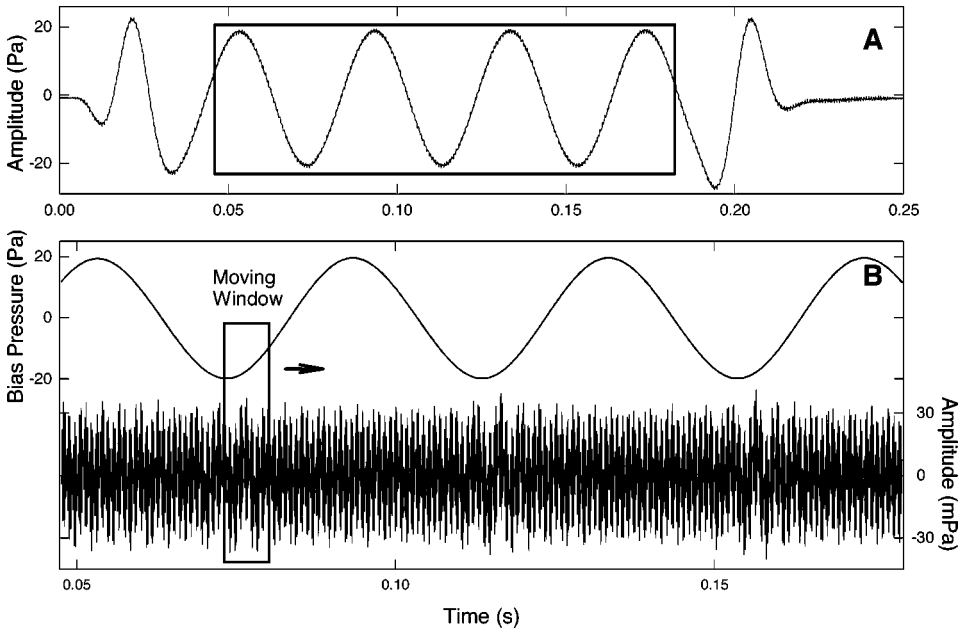


FIG. 2. (a) An example of the ear-canal acoustical signal. The peak amplitude of the bias tone is 120 dB pSPL (20 Pa) and the primary level ($L_1=L_2$) is 55 dB SPL. Data within the rectangular box containing three cycles of bias tone with stable peak-amplitudes are used to obtain the MPs of CDT magnitude. (b) Moving-window method. The acoustical signal in panel A is low- and high-pass filtered to yield the upper and lower traces, respectively. A 512-point moving-window is applied to both time-series to extract the CDT magnitude via FFT and the bias pressure with the mean value. The window is shifted 16 points to obtain the next pair of CDT magnitude and bias pressure, until reaching the ends of the waveforms.

SPL ($L_1=L_2$). The acoustic stimuli, two primaries and a bias tone, were generated from an array processor (Tucker-Davis Tech., TDT AP2), passed through a three-channel D-A converter (TDT DA3-4) controlled by three attenuators (TDT PA4). The two tones were delivered with two earphones (ER-3A) coupled to the ports on the ER-10B. The bias tone was amplified (Optimus SA-155) and guided from a sub-woofer (Boston Acoustics) to the speculum with a silicon tube.

Ear-canal acoustic signal [Fig. 2(a)] recorded by the ER-10B was low-pass filtered at 16 kHz, amplified 29 dB (Stewart VBF 10M), and digitized at 65 536-Hz sampling rate (TDT AD2). Depending on the primary levels, the signal was averaged 50 to 150 times with more averages at lower primary levels. The data was analyzed in MATLAB (v6.1, MathWorks) using a moving-window method (see below). The modulation pattern (MP) was obtained by plotting the CDT magnitudes against the instantaneous bias pressures within three cycles of the bias tone. Indices characterizing the MP of the CDT magnitude were measured separately for loading (increasing in bias pressure) and unloading (decreasing in bias pressure) processes from the data using an automatic program with interactive verifications from the experimenters. These indices obtained from the three cycles of the bias tone were then averaged independently for the two trials, two pairs of primary frequencies, and five primary levels. To assess effects of these experimental conditions, the physiologic indices of the CDT double MP were submitted to a two-way (trial×primary level) or a three-way (trial×loading/unloading×level) analysis of variance (ANOVA) in STATISTICA (v5.5, StatSoft). Data for the two pairs of primary frequencies were analyzed separately. A p-value of less than 0.01 was considered statistically significant.

C. Moving-window method

As illustrated in Fig. 2(b), the time waveform was first high-pass filtered at 200 Hz to obtain a time series consisting of the two primary tones and the DPs (lower trace). A window of 512 data points (7.8 ms) was taken from the time series starting from time zero. After Kaiser windowing (tenth-order), a CDT magnitude was extracted from the fast Fourier transformation (FFT). Then the window was shifted 16 points (0.24 ms) to obtain the next CDT magnitude until the window reached the end of the time waveform. Only the CDT data within the middle three cycles of bias tone, where the peak-amplitudes were stable (box in panel a), were selected for further analysis. The original data was also low-pass filtered at 100 Hz to obtain a signal containing only the bias tone [Fig. 2(b), upper trace]. The same 512-point window was taken from the bias tone data and the bias pressure was equal to the mean value within the window. This yielded a set of bias pressures and corresponding CDT magnitudes in the time domain.

D. Middle-ear experiments

Middle-ear transfer characteristics were measured from three gerbils to study the possible role of the middle ear in the formation of the MP of CDT magnitude. The animal preparation was similar to the DPOAE experiment, except the bulla was drilled open to gain access to the middle ear. The same bias tone (25 Hz, 120 pSPL) used in the DPOAE experiment was delivered to the ear. A laser beam (Laser Doppler Vibrometer HLV-1000, Polytec) was focused on the stapes footplate via an operating microscope (Zeiss 170) to obtain the middle-ear response. The sensitivity of the laser vibrometer was set at 1 mm/s/V. The stapes velocity (V_{ST}) was low-pass filtered at 16 kHz, amplified 5 dB (Stewart VBF 10 M), and digitized at a sampling frequency of 65 536 Hz (TDT AD2). The ear-canal acoustical responses were also recorded with the ER-10B microphone. Both signals were averaged 50 times. The middle-ear transfer function was obtained by plotting the V_{ST} versus the ear-canal sound pressures.

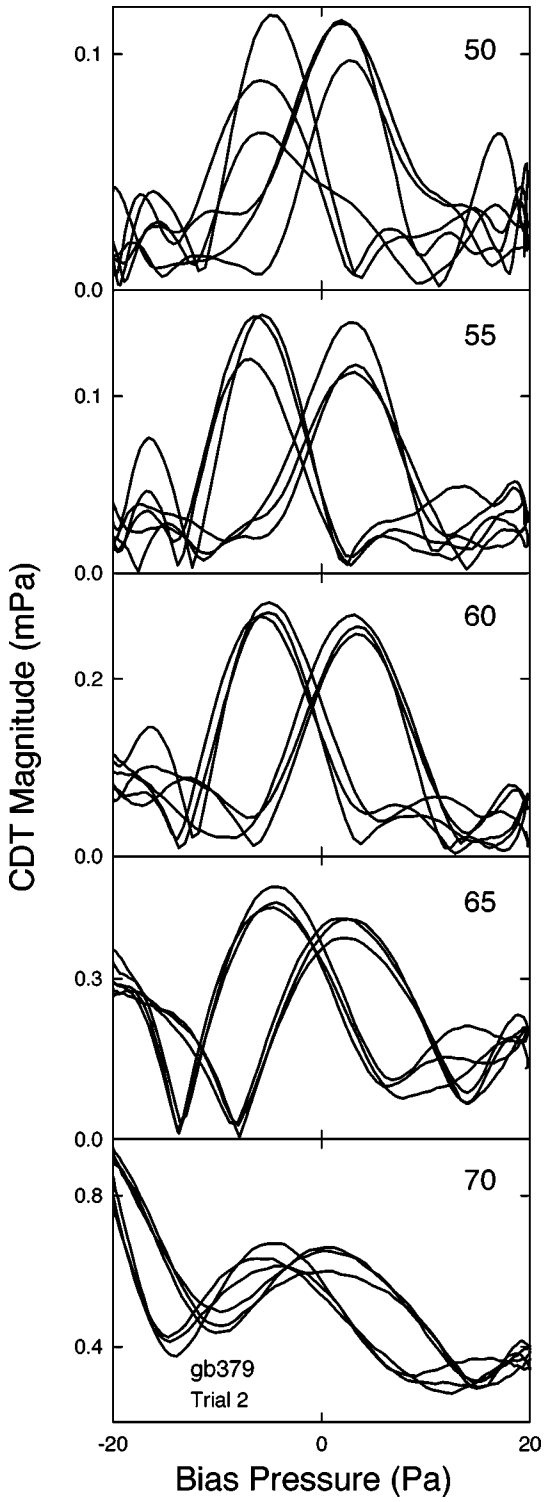


FIG. 3. Examples of the CDT double MPs. Indicated in each panel is the primary level in dB SPL ($L_1 = L_2$, $f_1 = 5888$ Hz, $f_2 = 7552$ Hz). The CDT data were obtained from three complete cycles of bias tone.

III. RESULTS

A. Modulation of CDT magnitude

At all the primary levels, the measured CDT magnitude as a function of the bias pressure showed a double MP containing two typical MPs (Fig. 3) similar to the shape of the $|f_{Tr}^{(3)}|$ which consists of a center lobe and two smaller side lobes [Fig. 1(d)]. In the time domain [Fig. 4(a)], this typical MP appeared for every half cycle of the bias tone with monotonic increase or decrease in amplitude. The time-domain CDT magnitude is consistent with the CDT wave envelope obtained with a narrow-band filter centered at the CDT frequency (Bian *et al.*, 2002). When plotted as a function of the bias pressure, the double MP of the CDT is marked by the separation of the center-peaks on the two typical MPs. The center-peak on the typical MP during loading of the cochlear transducer is located in the positive bias pressures [Fig. 4(b)], whereas the typical MP centered at negative bias pressures is related to the unloading process. As bias pressure varies with time, the two typical MPs with separate locations of the center lobes presented alternately [Figs. 4(a) and (b)]. This indicates that the maximal CDT magnitude occurs at a more positive or negative bias pressures depending on the direction of the bias pressure change or the loading/unloading of the cochlear transducer.

Since the f_{Tr} can be obtained by fitting the CDT-bias pressure data with a formula of the $|f_B^{(3)}|$ (Bian *et al.*, 2002), originally the two typical MPs related to loading or unloading processes were curve-fitted separately. This yielded two sigmoid-shaped functions with center portions separated from each other. Because the two typical MPs connected at the maximal bias tone amplitudes [Figs. 3 and 4(b)], the two sigmoid-shaped functions derived from curve fitting should be continuous at both extremes of the bias pressure, thus forming a hysteresis loop [Fig. 4(c)]. During loading of the cochlear transducer, the output follows the lower sigmoid-shaped curve whose center portion is in the positive bias pressures, whereas when it is unloading the transducer generates the upper curve with center located in the negative bias pressures. Therefore, the traversal of the cochlear transducer hysteresis loop is *counterclockwise*.

B. Characterization of CDT double MP and hysteresis

As illustrated in Fig. 5(a), the typical MP of the CDT magnitude was characterized with a set of physiological indices that have implications on the shape of the cochlear f_{Tr}. For the typical MP, the amplitude of the center lobe was measured as modulation height (MH) in dB SPL. The MH represents the CDT magnitude when the effect of the low-frequency biasing is minimal. The location of the center lobe, evaluated with the bias level at the peak, corresponds to the inflection point (IP) of the sigmoid-shaped f_{Tr} where the $f_{Tr}^{(2)}$ equals zero [Fig. 1(c)]. This bias pressure at the IP, denoted as BP_{IP}, also corresponds to the location where the slope or gain of the f_{Tr} is optimal since the $f_{Tr}^{(1)}$ is maximal [Figs. 1(a) and (b)]. The magnitudes of the left and right sidelobes are referred to as h_1 and h_2, also in dB SPL. Both h_1 and h_2 are the CDT magnitudes when the cochlear transducer is biased to the saturation ranges in negative or positive sound pressure directions, respectively. The asymmetry of the f_{Tr} is accessed by the ratio of h_1/h_2, with a value of one indicating perfect symmetry. The lower of the two notches on the typical MP, referred as index notch, corresponding to the maximal $f_{Tr}^{(2)}$ [Fig. 1(c)], was recorded to assess the lowest CDT magnitude due to the low-frequency biasing. A measure of the relative height of the center peak

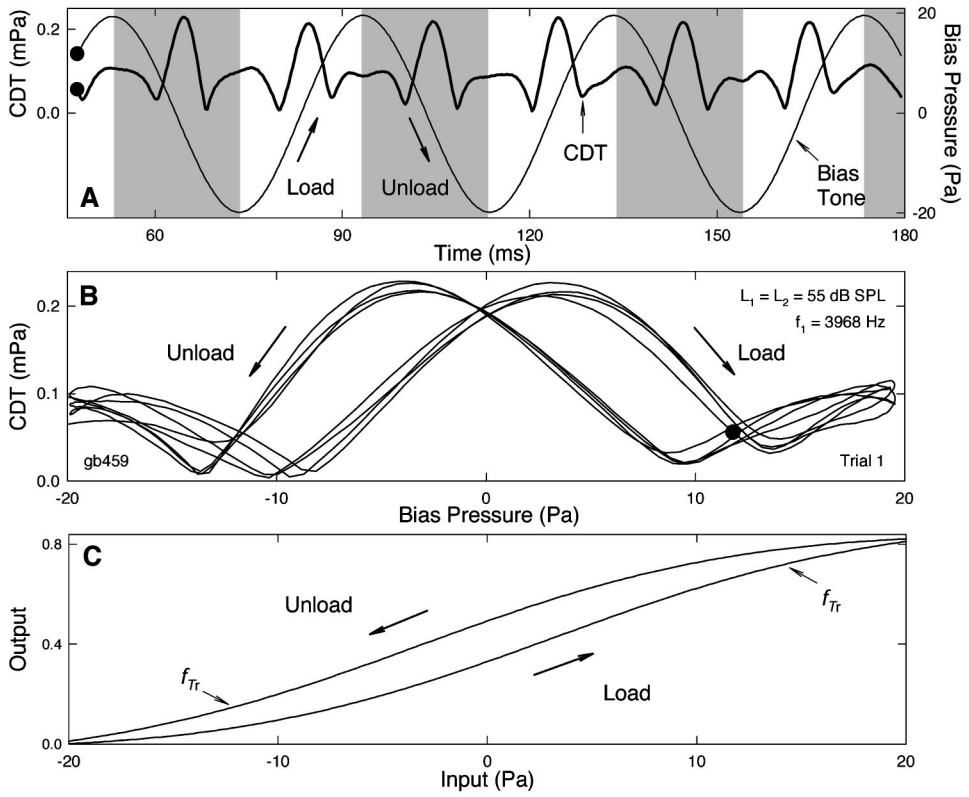


FIG. 4. (a) Low-frequency modulation of CDT magnitude in the time domain. The data were obtained from a segment of the waveform containing three cycles of bias tone with stable peak-amplitudes. Two similar MPs present alternatively as the sound pressure of the bias tone varies (increasing: loading; decreasing: unloading). (b) Double MP. The CDT magnitudes in panel A plotted as a function of the bias pressures show two separate MPs depending on the loading processes. Dots in panels A and B: starting points of the traces. (c) Hysteresis. Curve fitting the CDT-bias functions separately results in two sigmoid-shaped curves forming a *counterclockwise* hysteresis loop.

with respect to the lower notch was evaluated as modulation depth (MD) in dB. The MD reflects the maximal amount of reduction in CDT magnitude that is caused by the low-frequency bias tone. The distance between the two notches was measured with modulation width (MW). It simply indicates a portion of the transducer operating range where the transduction is nearly most efficient. For all the animals, the above physiologic indices were measured separately for both loading and unloading processes [Fig. 5(b)].

1. Separation of the center-peaks in the double MPs

Separation of the center-peaks of the CDT double MP was presented in all the animals for two pairs of primary frequencies at five primary levels. Since the center-peak of the MP corresponds to the IP of the sigmoid-shaped f_{Tr}, the separation of the two f_{Tr}, or the loop-width (LW) of the cochlear hysteresis, can be measured directly from the CDT data, i.e., the distance between the center-peaks of the double MP [Fig. 5(b)]. Because the statistical two-way ANOVA showed no significant difference between the repeated measures over 90 min, LWs obtained from the two trials were combined in estimating the mean and standard error (SE). Averaged across animals, a separation of about 8 to 9 Pa presented at all the primary levels for the two pairs of primary frequencies [Fig. 6(a)], and slightly reduced at higher primary levels. The separation of the two single MPs was also clearly demonstrated by the locations of the center-peaks or BP_{IP} [Fig. 6(b)]. At both pair of primary frequencies, the mean BP_{IP} across animals and trials (ANOVA, $p > 0.01$) for loading is about 3 Pa, whereas the BP_{IP} of the unloading process is located at about -6 Pa. For both loading and unloading processes, the BP_{IP} shifted towards negative sound pressures at the highest primary level, 70 dB SPL. The double MP (Figs. 3 and 4) is the major effect of the low-frequency biasing on the CDT magnitude.

2. Effects of primary level on double MPs

Other physiological indices of the MP measured for both loading and unloading processes at either primary frequencies showed significant effects of the primary level (Fig. 7). Since the three-way ANOVA showed no statistical significant difference presented between the two trials, the data were pooled to calculate the mean and SE for each variable. Indices MH, notch, h_1, and h_2 plotted as functions of primary levels showed different growth patterns. Panel (a) of Fig. 7 compares the growths of CDT magnitudes when the modulation is minimal (MH) and maximal (notch). The MH growth function is generally higher in magnitude but with a shallower slope (about 1 dB/dB) and dynamic range. However, CDT magnitude at the notch is much smaller and rises at a greater, 2 dB/dB, rate. As shown in panels (b) and (c), the growths of h_1 and h_2 fall between the growth functions of MH and notch. At higher primary frequencies ($f_1 = 5888$ Hz), the MH, h_1, and h_2 are slightly larger than those of the lower primaries ($f_1 = 3968$ Hz), especially around 60 dB SPL primary level. However, there is no such difference in the growth functions of the notch. The growth patterns of these indices were similar for both loading and unloading processes (ANOVA, $p > 0.01$).

Reduction in CDT magnitude due to the bias tone was greater at low primary levels as indicated by the index MD [Fig. 7(d)]. There were about 25-dB reductions in CDT magnitude at 50 or 55 dB SPL primary level compared to only 5-dB modulation at 70 dB SPL. In other words, low-

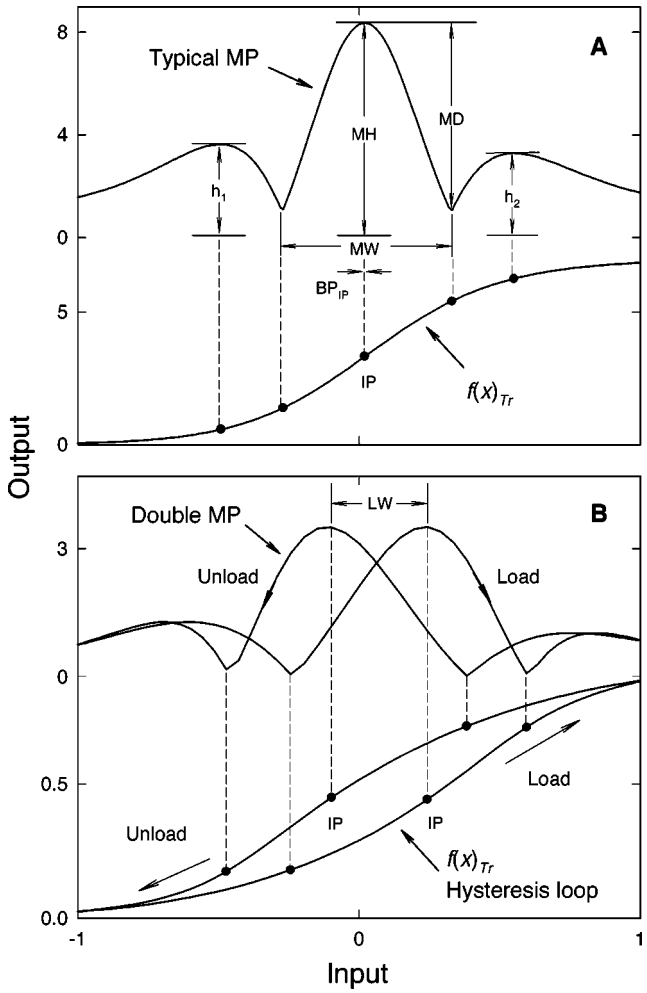


FIG. 5. Characterization of the MP of CDT magnitude. (a) Typical MP and f_{Tr}. The typical MP contains a center lobe and two sidelobes. MH, h_1, and h_2 are the amplitudes of the center- and sidelobes, respectively. MD: relative height of the center lobe to the notch. MW: width of the center lobe is measured by the distance between the two notches. Corresponding locations of each of the indices on the f_{Tr} also shown in the lower curve. (b) Double MP and f_{Tr} hysteresis. The separation of the two center-peaks on the double MP is used to determine the loop width (LW) of the hysteresis. Corresponding locations of the center peaks and notches of the MP on the hysteresis loop are indicated.

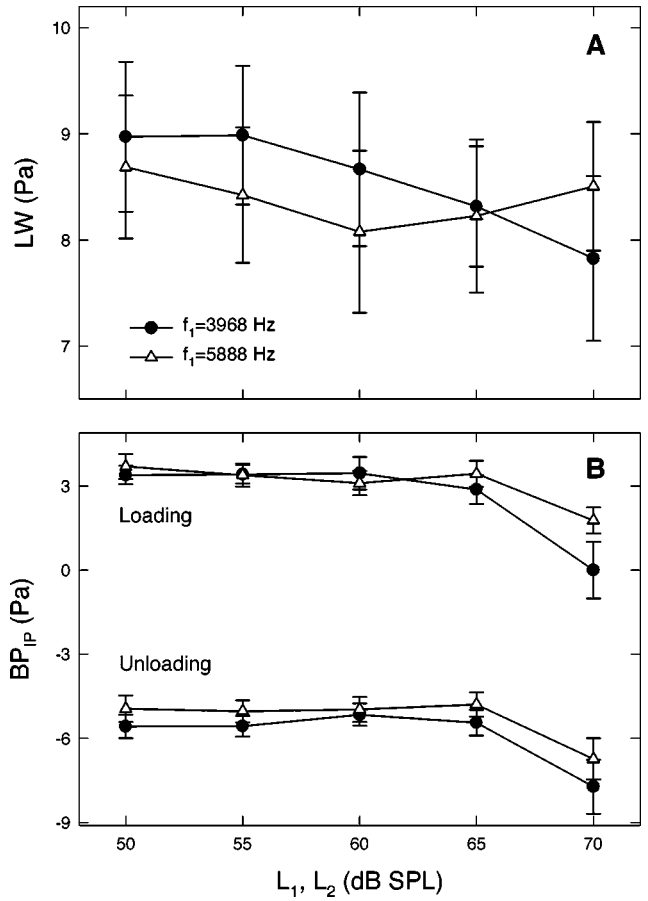


FIG. 6. Separation of the center-peaks on the double MP of CDT magnitude. Mean±standard error (SE) of the LW (panel a) and BP_{IP} (panel b) for the two pairs of primary frequencies. Data reflect the average across two trials ($n=20$).

frequency modulation was more effective at low primary levels. In terms of the MP, the notches were deeper at low signal levels, and shallower at high level. Panel (e) of Fig. 7 shows that the MW expands from about 15 to 30 Pa as primary level grows from 50 to 70 dB SPL. This indicates that the two notches of the MP are further apart when primary level increases. The mean of the asymmetry index, ratio of h_1/h_2, was greater than one at all the primary levels [Fig. 7(f)]. This implies that the left sidelobe of the MP located in the negative bias pressures is higher than the right one on the positive pressure side. Moreover, the asymmetry increased with primary levels and decreased with primary frequency.

C. Middle ear transfer characteristics

To investigate the possible contribution of the middle ear to the CDT double MP and hysteresis, vibrations of the stapes in response to the low frequency (25 Hz) bias tone were measured with laser vibrometry from three gerbils in a separate experiment. Figure 8(a) shows an example of the middle ear V_{ST} (containing high frequency noise) and the ear-canal sound pressures (P_E). Since the middle-ear response is a velocity measure, it has a 90°-phase difference with the acoustic signal. Thus, when the acoustical signal or P_E was shifted 10 ms or $\frac{1}{4}$ of the period, it overlapped with the trace of the V_{ST}. The middle-ear transfer function then was obtained by plotting the V_{ST} against the phase-shifted P_E (panel c). This middle-ear transfer function demonstrated essentially a linear relation between the acoustic signal and the stapes response. Such relation can be better visualized by low-pass filtering the V_{ST} at 200 Hz to eliminate the high-frequency noise in the raw data (panel b). The filtered V_{ST} and the phase-shifted P_E overlapped with each other very well, indicating a linear middle ear transfer characteristic without hysteresis [Fig. 8(d)]. Three animals all showed such linear middle ear transfer function for sinusoids with a peak amplitude of 20 Pa.

This result can exclude the possible involvement of the middle ear in the formation of the double MP of CDT magnitude and the associated hysteresis loop. The linearity of the middle-ear transmission is consistent with other observations in gerbils (Dirckx *et al.*, 1998; Lee and Rosowski, 2001) that the tympanic membrane movement starts to saturate beyond 200 Pa of input pressure. In these studies, hysteresis could only be observed with input pressure well above the sound

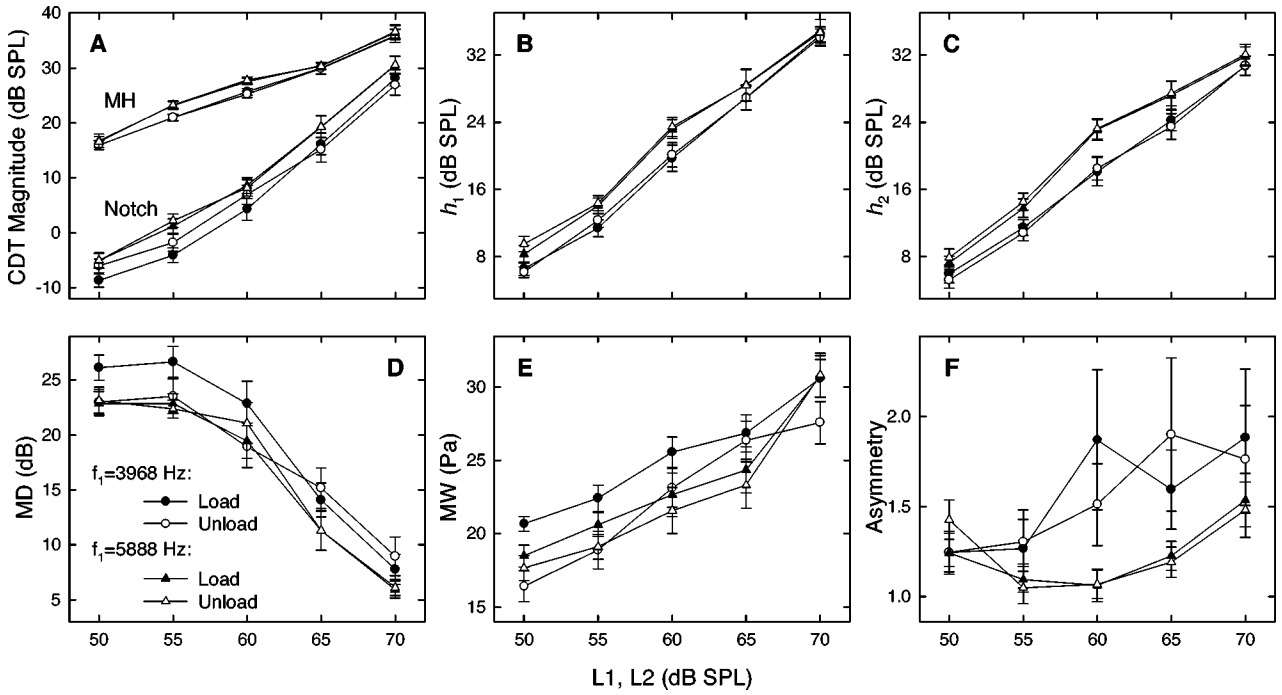


FIG. 7. The effect of primary level on physiologic indices. (a–c) Different growth patterns for MH, notch, h_1, and h_2. (d) MD reduces with primary level. (e) MW expands with level. (f) Asymmetry index is greater than one for all primary levels. The MP becomes more asymmetrical towards higher primary levels and at low frequencies. Data represent mean±SE of the indices averaged across the two trials ($n=20$). These effects are consistent among the loading and unloading processes.

pressure used in the present study (20 Pa). A recent study on gerbils (Choi *et al.*, 2002) also showed linear middle-ear transfer functions obtained with a wideband noise. Stapes vibrations measured from neonate and adult gerbils (Overstreet and Ruggero, 2002) support the assumption that the middle ear works as a transmission line to deliver the acoustic signal to the inner ear without distortion and dissipation. Therefore, the nonlinearity and hysteresis observed with the low-frequency modulation of CDT originate from the inner ear.

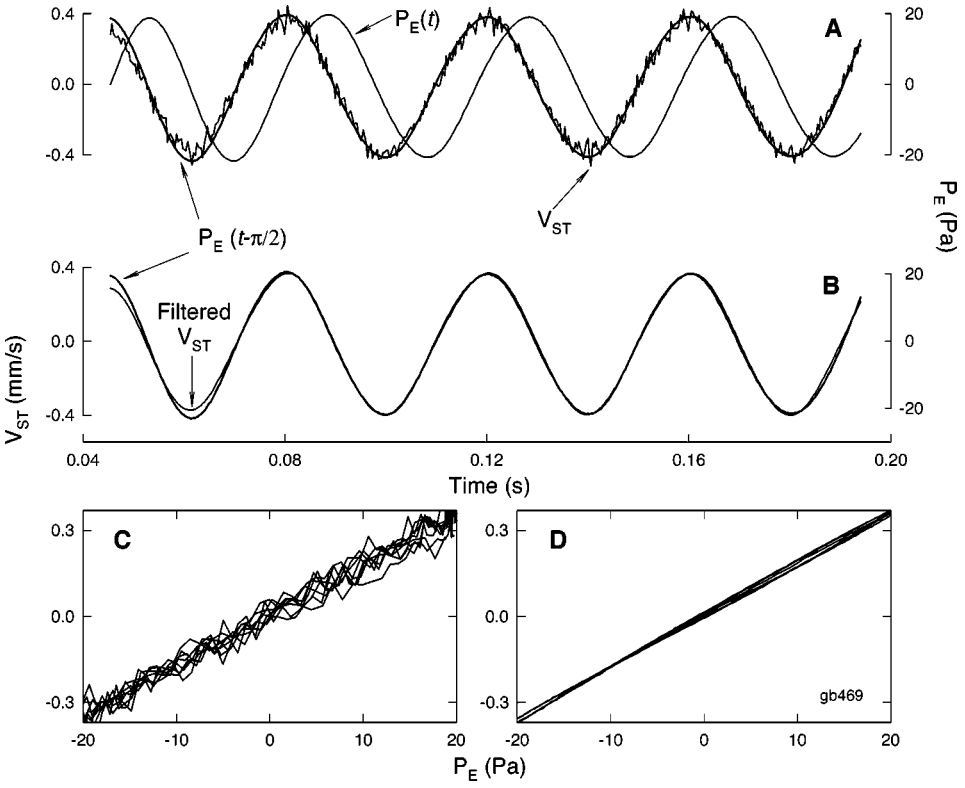


FIG. 8. Obtaining the middle ear transfer characteristics. (a) Ear canal sound pressure (P_E) and the corresponding stapes velocity (V_{ST}). Note: there is a 90°-phase difference between V_{ST} and P_E. When P_E is shifted by $\pi/2$ or 1/4 cycle, it overlaps with the V_{ST}. (b) Low-pass filtered V_{ST} vs. phase-shifted P_E. To eliminate the high-frequency noise, the stapes response was low-pass filtered at 200 Hz. (c) Middle ear transfer function: original V_{ST} (containing high frequency noise) vs. phase-shifted P_E. (d) Middle ear transfer function: low-pass filtered V_{ST} vs. phase-shifted P_E.

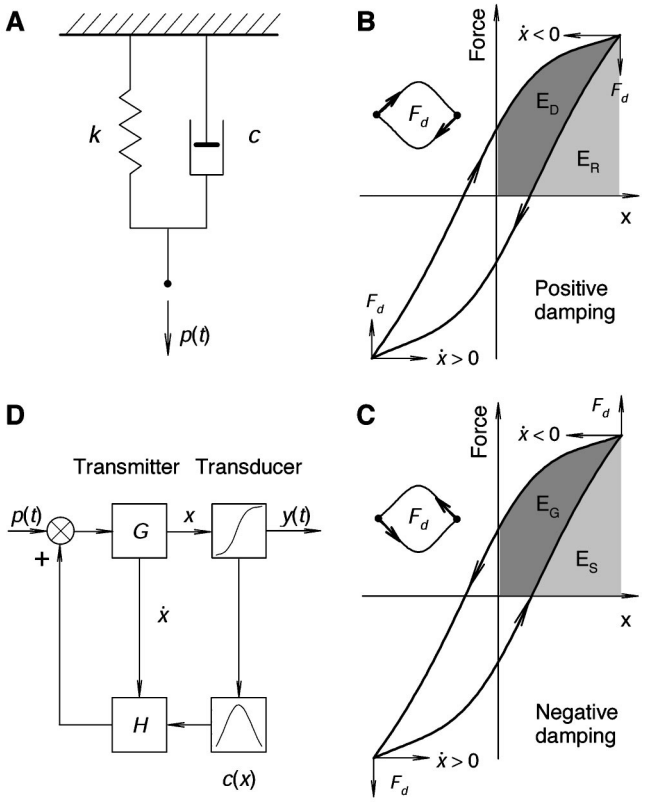


FIG. 9. Model for cochlear mechanical hysteresis. (a) A Kelvin-Voigt spring (k)-damper (c) system with an external excitation $p(t)$. A sigmoid-shaped force-displacement relation is assumed for the k. (b) Effect of positive damping on the force-displacement relation: *clockwise* hysteresis. The damping force (F_d) always has the same sign as $\dot{x}$. Shaded areas represent the energy (E) transformation during a half cycle of motion ($x>0$). $E_R: E_{released}$; $E_D: E_{dissipated}$. (c) Effect of negative damping: *counterclockwise* hysteresis. The F_d always has the opposite sign of $\dot{x}$. $E_S: E_{stored}$; $E_G: E_{gained}$. (d) Block diagram of the feedback model. The $p(t)$ passes through a transmitter with gain of G, then becomes the input x to the transducer. During the transduction process, the output $y(t)$ is generated, and the F_d with a gain H is fed back to the input. The traversal of the hysteresis loop generated by this model depends on the sign of $c(x)$.

IV. MODEL

A. A model for cochlear hysteresis

A striking feature of the hysteresis loop of the f_{Tr} is the counterclockwise direction of traversal [Fig. 4(c)]. Since the DPOAEs are mechanical responses, the derived f_{Tr} represents a force-displacement relation of the cochlear transducer, where the lower curve of the loop is the effect of loading and the upper curve the unloading effect. For most mechanical systems, the typical traversal of the stress-strain hysteresis is clockwise, indicating energy dissipation within the system (Lakes, 1999; Tongue, 2002). Analogously, counterclockwise traversal of the hysteresis suggests that the cochlear transducer produces energy during the sinusoidal loading processes. A one-dimensional Kelvin–Voigt viscoelastic model of the cochlear transducer containing parallel spring and damper could help to identify the source of the energy gain [Fig. 9(a)]. For viscoelastic systems under sinusoidal loading, the mechanical response consists of in- and out-of-phase components that are determined by a complex modulus. Dynamic stiffness, $k(x,t)$, the real part of the modulus, reflects the energy conversion which is in phase with the loading process. As indicated by experiments on isolated hair cells (Hudspeth, 1989; Evans *et al.*, 1991; He and Dallos, 1999), the force mainly due to k follows a sigmoid-shaped f_B. The imaginary component, the damping coefficient $c(x,t)$, governs the energy loss or gain within the system and the associated phase shift in the response. Adding a damping force ($F_d=c\cdot\dot{x}$, $\dot{x}$ is velocity) can modify the force-displacement relation depending on the sign of c and the direction of motion, since the total force (F) constitutes stiffness force (governed by the f_B) and F_d, or

$$F(x)=f_B(x)+F_d(x). \tag{2}$$

For positive c [Fig. 9(b)], the sign of the F_d is always the same as $\dot{x}$, resulting in greater force during loading and smaller force for unloading, i.e., a *clockwise* hysteresis. However, in a system with negative c where F_d is proportional to $\dot{x}$ with opposite sign [Fig. 9(c)], the traversal of the hysteresis turns *counterclockwise* (Peters and Pritchett, 1997; Lakes, 1999). The energy transformation in the passive and active systems can be illustrated with the area under the force-displacement curve. This area indicates potential energy stored (E_S) when the direction of motion is away from its equilibrium ($x=0$), and kinetic energy released (E_R) when it is returning to zero displacement. As shown in Figs. 9(b) and (c), for a passive system ($c>0$), the E_S is always greater than the E_R, since $E_S=E_R+E_{dissipated}(E_D)$; and in an active system ($c<0$) more energy is released than stored, i.e., $E_R=E_S+E_{gained}(E_G)$. If $c=0$, the hysteresis loop is reduced to a single sigmoid-shaped function, because $E_S=E_R$.

B. Fit to CDT data

To simulate the experimental data, the model in Eq. (2) was realized with a local-memory feedback mechanism [Fig. 9(d)], where F_d serves as the feedback component to the f_B. Since the exact BM displacement could not be determined, a relative damping factor $C(x,t)$ that can be approximated from the $f_B^{(1)}$ [Fig. 1(b)] was adopted to evaluate the F_d. One reason for such approximation came from the observations that the dynamic $k(x)$ of stereocilia or OHC body is voltage- or displacement-dependent due to the Boltzmann process in MET (Hudspeth, 1989; He and Dallos, 1999). Another reason was that the hysteresis loop is widest around the IPs of the f_{Tr} (Figs. 4 and 5), where energy exchange is more efficient and the $f_{Tr}^{(1)}$ is maximal. In the model, the ear-canal input sound pressure (p) was matched to the input impedance of the cochlea transducer (Patuzzi, 1996) with a gain G due to the transmission and distribution of sound energy on the BM. This instantaneous input to the transducer produced a force $y(x)_t$ and a $C(x)_t$ via f_B and $f_B^{(1)}$ (Bian *et al.*, 2002). Then the F_d was fed back to the input with a gain of H to determine the force generation at the next moment in time. Thus, the feedback mechanism could be encapsulated in the f_B of Eq. (1) to compute the force at time t, i.e.,

$$F(x)_t=f_B[G\cdot x_t+H\cdot F_d(x)_t], \tag{3}$$

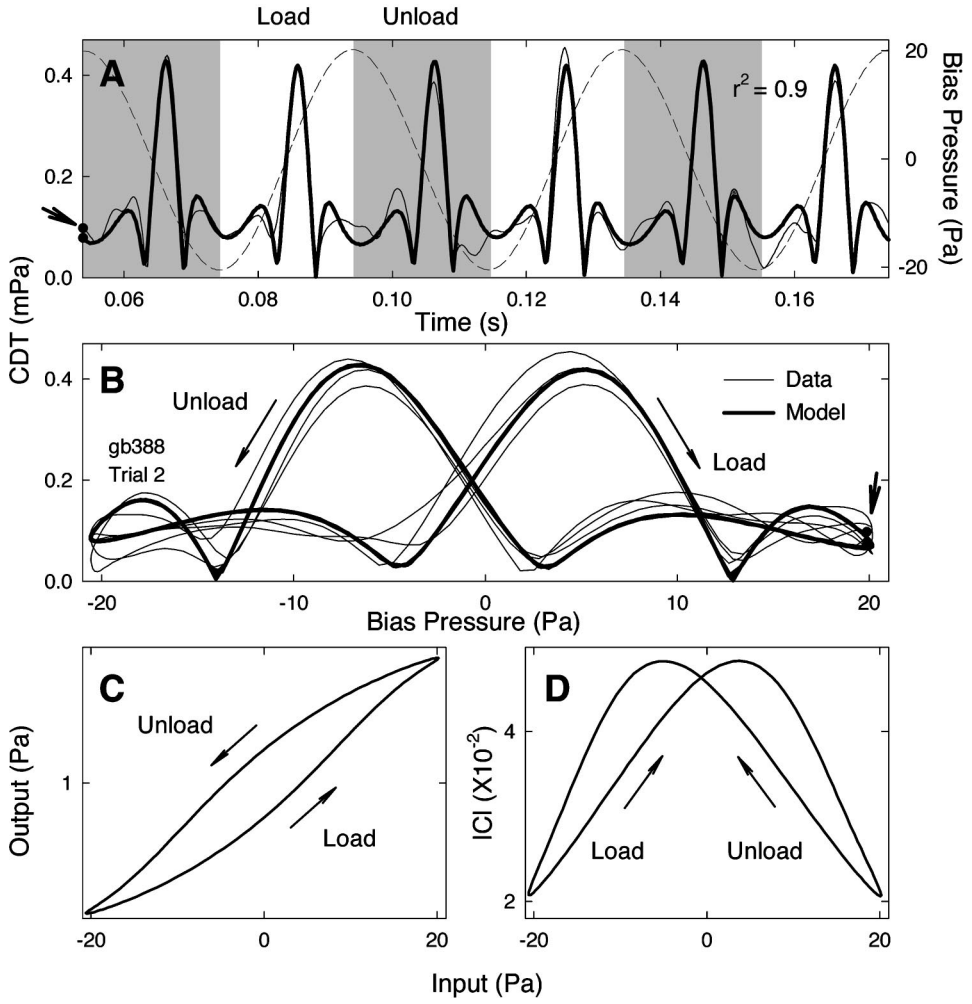


FIG. 10. Model results. (a) CDT magnitude produced by the model (thick line) vs. experimental data (thin line) in the time domain. Dashed line is the bias tone. (b) Model vs. data: CDT double MP. The CDT magnitudes in panel (a) are plotted as a function of bias levels. The short arrows indicate the starting points of the traces. Data were collected from a two-tone signal at 55 dB SPL (f_1=5888 Hz, and f_2=7552 Hz). (c,d): The hysteresis loop and the absolute value of the damping factor $C(x)$ in arbitrary unit obtained from the model for the data in panels (a) and (b).

where $F_d(x)_t = c(x)_{t-\tau} \cdot \dot{x}_t$, and $\dot{x}_t$ is the instantaneous velocity of motion, i.e., $(x_t - x_{t-\tau})/\tau$, where τ is the time increment of the waveform.

The same bias tone and two-tone signal used in the experiment [Fig. 2(a)] were sent to the model. With predetermined initial values for the model parameters, the output was produced in the time domain using an algorithm described in Eq. (3). The moving-window method was applied to the output to extract the CDT magnitude and bias level. A correlation coefficient (r^2) and a sum of squared error were calculated between the fit and the experimental data [Figs. 10(a) and (b)]. A function minimization routine (Nelder-Mead simplex method, MATLAB) was used to find the model parameters that had the least sum of squared errors and maximal r^2. The initial values of the parameters were varied and the fitting process repeated to avoid local minima in the results. The model was fit to each animal for all the experimental conditions. The resulting model parameters were subjected to a two-way (trial×primary level) ANOVA.

As shown in panel (a) of Fig. 10, the model produced modulation of CDT magnitude in the time domain that matched the CDT data. The simulated CDT magnitude as a function of bias pressure showed a double MP similar to the experimental results [Fig. 10(b)]. The cochlear f_{Tr} and the absolute value of C derived from the model demonstrated hysteresis and peak-separation [Figs. 10(c) and (d)]. Like the f_{Tr} obtained from the experimental CDT data, the traversal of the modeled hysteresis loop was counterclockwise. The factor C shifted to the positive sound pressures while loading the cochlear transducer, and moved to the negative direction during unloading. Since the ANOVA showed no significant difference between the two trials, the model parameters were averaged across trials and summarized in Table I and Fig. 11. For the three lower primary levels (50–60 dB SPL) that show the most typical CDT double MP, 55 dB SPL is the level where the mean r^2 is among the highest.

Most of these parameters, namely A, H, G, b, c, and d, demonstrated significant effects of the primary level. Table I lists the model parameters that are equivalent to the constants in the second-order f_B [Eq. (1)]. Parameter A, the maximal mechanical response, increased about 30 to 60 times or 30 to 35 dB over a 20-dB range of the primary level. Parameters b and d are the slope constants that may reflect the gain of the cochlear transducer. The absolute values of b and d reduced three to four times or 9–12 dB when primary level increased from 50 to 70 dB SPL. Parameters c and e relate to the position of the transducer curve. Evaluated with the bias pressure, both c and e showed a negative-direction shift with increasing in primary level. In other words, the f_{Tr} shifted to the negative sound pressures at higher primary levels. Two dimensionless constants, H and G, showed consistency among the animals and between the two pairs of primary

TABLE I. Model results: the mean and SE of the Boltzmann parameters of a model of cochlear hysteresis which was fit to the CDT data and the correlation coefficients (r^2) of the fits for different primary frequencies and levels. The means reflect an average of the two trials.

f_1 (Hz)	L_1, L_2[a] (dB SPL)	A	b[b]	c	d[b]	e	r^2
3968	50	0.47±0.07	−9.5±0.67	−3.2±0.91	5.9 ±0.90	12.0± 3.4	0.62±0.05
	55	2.2 ±0.44	−8.2±0.68	−9.1±1.3	2.1 ±0.81	4.9± 6.3	0.76±0.03
	60	4.0 ±0.64	−8.0±0.79	−7.2±2.3	3.4 ±0.84	−5.6± 3.6	0.72±0.05
	65	13.7 ±4.5	−4.6±0.88	−11.9±5.7	1.2 ±1.1	−4.1± 4.3	0.83±0.03
	70	30.3 ±6.3	−3.2±0.49	−15.5±4.9	1.4 ±0.51	−2.7± 6.2	0.91±0.01
5888	50	0.52±0.05	−9.0±0.68	−7.1±2.4	2.8 ±1.4	5.2± 7.5	0.65±0.04
	55	1.4 ±0.22	−8.8±0.63	−2.3±3.3	2.4 ±1.1	9.0±10.3	0.83±0.02
	60	4.3 ±2.1	−7.8±0.85	−11.0±2.6	1.6 ±0.73	−9.2±13.6	0.70±0.07
	65	5.9 ±1.4	−6.7±1.8	−20.3±6.0	2.4 ±1.6	−1.7±16.9	0.64±0.07
	70	17.7 ±5.0	−3.1±0.37	−27.1±2.3	0.87±0.17	−6.2±42.0	0.81±0.14

Mean±SE.
[a]$L_1 = L_2$.
[b]$\times 10^{-2}$.

frequencies (Fig. 11). The feedback gain H was larger at lower signal levels, and reduced exponentially (about 20 dB) with increasing in the primary level (panel a). Examining across primary levels, the H was greater for the higher primary frequency ($f_1 = 5888$ Hz). Finally, G, gain from the impedance matching, showed a similar pattern with a much smaller magnitude [Fig. 11(b)]. At low primary levels, the G was greater for the lower primary frequency ($f_1 = 3968$ Hz). These effects indicate that the input impedance of the cochlear transducer is smaller at low signal levels and low frequencies.

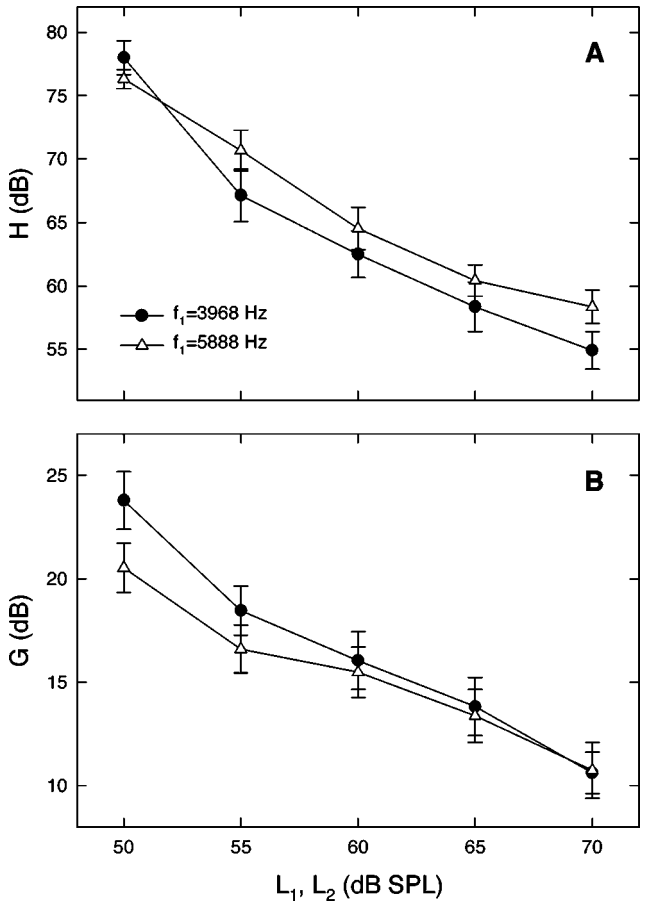


FIG. 11. The mean±SE of the feedback gain H (panel a) and the transmitter gain G (panel b) for the two pairs of primary frequencies ($n = 20$, two trials combined).

V. DISCUSSION

A. CDT double MP

The key findings of the present study are the double MP of CDT magnitude and the associated cochlear hysteresis. Low-frequency modulation of DPOAEs has been observed, but the double MP and hysteresis have never been reported. In a study in humans on low-frequency biasing of DPOAEs by Scholz *et al.* (1999), the modulated CDT magnitude in the time domain showed periodicity over one cycle of bias tone marked by two peaks. These two maximal CDT amplitudes could correlate with the center lobes of the double MP found in the present study even though their data did not show the same MP. Thus, the double MP of DPOAEs and hysteresis may be present in humans. Besides species difference, lower bias tone amplitude may contribute to the different MPs observed by Scholz *et al.* (1999). In gerbils, Frank and Kössl (1997) noticed that the maximal CDT magnitude occurred at the zero-crossings of the bias tone, which is in accordance with the double MP during one cycle of the bias tone observed in the present experiment. Electrically evoked otoacoustic emissions (EEOAEs) can be modulated by a low-frequency bias tone (Kirk and Yates, 1998) and show an enhanced peak when the BM was biased in one direction. Moreover, Kirk (2002) observed a double MP of EEOAEs after application of ATP in scala media (SM), i.e., two ''enhanced'' EEOAEs wave envelopes appeared within a single period of bias tone. This may indicate that the hysteresis in cochlear transduction responsible for the generation of OAEs requires energy or ''active.''

The center-peaks occur at a slightly higher bias pressure during loading and more negative pressures during unloading. In time domain, this indicates that the CDT response is out-of-phase relative to the bias tone, or delayed. Indeed, when the CDT waveform is shifted *backward* in time, the separation of the two center peaks can be diminished. Thus

the time delay can be estimated as the number points shifted to completely overlap the two center-peaks multiplied by the time increment, i.e., 0.24 ms. Averaged across animals and primary levels, this time delay ranged from about 1.4 to 1.2 ms for the two pairs of primary frequencies (3968/5120 and 5888/7552 Hz), respectively. These time delays are consistent with the CDT group delay measured in gerbils (Mills and Rubel, 1997; Faulstich and Kössl, 2000) and onset latency in rabbits (Whitehead *et al.*, 1996). Estimated with the cochlear microphonic (CM) recorded along the cochlear partition, the one-way propagation time to the CF place of 5 kHz is about 0.1 ms (Schmiedt and Zwislocki, 1977). The CDT time delay is considerably longer than the ''round-trip'' propagation time of the traveling wave. Therefore, the time delay is mainly due to the activation of the bi-directional transduction processes, possibly by a time-dependent feedback mechanism. Direct measurements of the BM vibration at CF place during ''low-side'' suppression show two phases of CF suppression within one cycle of the low-frequency tone and time delay of the CF component (Cooper, 1996; Geisler and Nuttall, 1997). The latter authors demonstrate a double MP similar to that of the *C* factor shown in panel (d) of Fig. 10 when plotting the envelope of the CF response (gain of the cochlear transducer) as a function of the BM displacement due to the suppressor. These time- or phase-differences between the CF/CDT and the low-frequency bias tone may also reflect the filter effect of the cellular or molecular processes within the cochlear transducer. Both the transduction channel gating and the OHC membrane motor motility are time-dependent and require finite response time or have frequency limit (Kros *et al.*, 1992; Gale and Ashmore, 1997). However, in future study, the exact propagation time should be determined to accurately measure the LW.

B. Effects of primary level

The MPs found in the present study show effects of the primary level that are consistent with our previous work (Bian *et al.*, 2002). As primary level increases, peaks and notches of the double MP demonstrate different growth patterns (top panels of Fig. 7) that are also indicated in the 3-D surface plot of the MPs in the previous paper. These growth patterns reflect shifts of the operating point (OP), a point on the f_{Tr} where the transfer characteristic operates on the input signal. Thus, the CDT magnitude is determined by the shape and slope of the f_{Tr} at the OP (Lukashkin and Russell, 1999). In the case of MH, higher magnitudes at lower primary levels indicate steeper slopes at the IPs where the OPs are located. Notches of the MP represent the transitions of the f_{Tr} from growth to saturation (Fig. 1). At lower primary levels, deeper notches relative to the center-peak could imply more abrupt slope changes at these transitions. These alterations of the f_{Tr} at the OP contribute to the variations in the CDT growth patterns observed in humans (Dorn *et al.*, 2001) and animals (Mills, 2002).

There is marked asymmetry in the double MP, especially at higher primary levels. Asymmetry index greater than one indicates that biasing towards negative sound pressures produces a higher CDT magnitude (h_1). Moreover, the IP with optimal gain shifted to the negative pressures at 70 dB SPL [Fig. 6(b)]. These asymmetrical aspects of the f_{Tr} could explain the observations of enhancement of EEOAEs while presenting a low-frequency tone (Kirk and Yates, 1998; Kirk, 2002) or a tone near the primary frequency (Ren and Nuttall, 1998; Nakajima *et al.*, 2000). The electrical currents used to evoke EEOAEs may be comparable to acoustic primary tones at high levels ($>$70 dB SPL) where the IP is located in the negative sound pressure range. Thus, low-frequency biasing in the negative direction, which corresponds to a shift of the BM towards scala vestibuli (SV), results in greater CDT magnitude. High-level tone presented near the primary-frequency locations could produce a static movement of the BM towards SV due to shortening of the OHCs (Zimmermann and Fermin, 1996). Indeed, a similar shift of the BM and a greater magnitude of vibration were observed while injecting positive current into the SM (Parthasarathi, 2003). This mechanism, asymmetry in f_{Tr} and shift in IP, could offer an explanation for a more complicated phenomenon: suppression and enhancement of two-tone DPOAEs while adding a third high-frequency tone (Mills, 2000; Martin *et al.*, 2003). Due to the shift of the f_{Tr} in the negative sound pressure direction at high signal levels, the CDT may be generated from the saturation region corresponding to one of the sidelobes on the double MP. This could partly explain the different vulnerability of the low- versus high-level CDTs to cochlear damages (Avan *et al.*, 2003). The slope of the f_{Tr} in the saturating region is relatively small and may not be further reduced; however, the optimal transducer gain at low signal levels is high and any cochlear insult could result in large reductions.

These level-dependent variations in the shape of the f_{Tr} are in agreement with the MET characteristics obtained with CM in response to broadband noise (Bian and Chertoff, 1998, 2001) and pure tones (Patuzzi and Moleirinho, 1998). Similarities in the transfer functions derived from DPOAEs and CM indicate that the bi-directional transduction processes are highly coupled, possibly via a positive feedback (Yates, 1990; Lukashkin and Russell, 1999). It is proposed that the OHC receptor current is the feedback signal that controls the cochlear amplifier gain with the shift of the OP (Kössl and Russell, 1992). Level-dependent alterations in the transfer characteristics suggest that the cochlear f_{Tr} dynamically change its shape to accommodate the input signal, or the cochlear transducer adjusts its optimal gain based on the intensity of the incoming sound.

C. Physiologic indices versus model parameters

Characteristics of the cochlear f_{Tr} can be directly measured from the CDT double MPs at various primary levels, and indirectly from fitting the CDT-bias level data using an f_B-based model with feedback. The Boltzmann parameters (Table I) derived from the model-fit are comparable to those reported in our previous work (Bian *et al.*, 2002). The signs of these parameters generally agree, and the values show more consistency in the present study. Model parameters in Table I demonstrate effects that are similar to some physiologic indices derived from the CDT double MPs. *A*, the maximal mechanical response of the transducer, grows exponentially at a nearly 2 dB/dB rate, similar to the growth of

notch and h_1 or h_2 on the MP. The notch and sidelobes relate to the saturation of the f_{Tr}, whereas A may refer to the mechanical responses without cochlear feedback. Two model parameters, b and d, represent the slopes of the upper and lower branches of the f_B with respect to the IP. They may be similar to index MD, a measure of the relative slope change from optimal to saturation. Both the slope parameters and MD show greater values at low primary levels, indicating higher optimal gain for detecting soft sound, a feature of ''cochlear amplifier'' (Gold, 1948; Davis, 1983). Model parameters c and e govern the location of the sigmoid-shaped f_B. Evaluated with the bias level, they show a trend of shifting to negative sound pressures as primary level increases similar to BP_{IP} [(Fig. 6(b)].

Index LW directly reflects the size of the hysteresis loop, which is unique for the method of direct measurement of the MP. The hysteresis loop is slightly wider at lower signal levels as indicated in Fig. 6(a). Two parameters, only available for the model-fit method, are the feedback gain H and impedance matching gain G. Parameter H is a gain from which the F_d can contribute to the total force generation. Higher H at low primary levels [Fig. 11(a)] indicates that the feedback mechanism is more important for amplification and detection of soft sound (Gold, 1948; Davis, 1983). Greater change in H with primary level than the LW may indicate that the H contributes not only to the formation of the hysteresis, but also to the relatively higher CDT magnitude at low primary levels. Greater G, gain from impedance matching, at low signal levels [Fig. 11(b)] may suggest that the size of the cochlear transducer perturbed by soft sound is smaller with lower impedance. Higher G at the lower primary frequency ($f_1 = 3968$ Hz) is consistent with the calculation of Overstreet and Ruggero (2002), who showed low cochlear input impedance at around 4 kHz and increasing with frequency. This difference may reflect the variation in the mechanical properties of the transducers along the cochlear partition.

Reliable estimates of cochlear f_{Tr} using the physiologic indices and model parameters should be limited to lower primary levels, because the accuracy reduces with increasing in signal level. Possible factors that influence the accuracy of the method include theoretical concerns, less effective modulation, and interference from other DPOAEs generation mechanisms or sites within the cochlea. As discussed previously (Bian *et al.*, 2002), the contribution from higher-order terms in the power series that approximates the nonlinear f_{Tr} could become significant at high primary levels, thus using the third derivative approach to derive the cochlear f_{Tr} may not be appropriate. The low-frequency modulation of DPOAEs may be less effective at high primary levels, simply because the biasing displacement of the BM is smaller compared to the response of the primaries. Distortions produced from multiple reflections within the cochlea (Shera and Guinan, 1999) could become prominent at high primary levels and may not be affected by the bias tone, thus resulting in elevated notches in the MP. However, if the CDT generated from the reflections in the cochlea is time-invariant, then it could be eliminated by shifting the MP downward until the notches reach zero.

D. Cochlear hysteresis and negative damping

Hysteresis has been observed in several different cochlear responses. Intracellular receptor potentials of both IHCs and OHCs show hysteresis loop when plotted as a function of the sinusoidal input (Cody and Russell, 1992). Extracellular CM measured in the scala tympani is hysteretic and level dependent (Patuzzi and Moleirinho, 1998). Cochlear MET function, obtained from the CM recorded at the round window and V_{ST} in the middle ear, demonstrates sigmoid-shaped hysteresis (Chertoff *et al.*, 2001). Measurements on the organ of Corti vibrations suggest that for the same input the resulting mechanical response can be different depending on the direction of motion (Hao and Khanna, 2000). Moreover, double MP of the BM vibration at the CF is observed in two-tone suppression experiments (Geisler and Nuttall, 1997). The envelope of the CF response as a function of the low-frequency displacement reveals the first derivative of a sigmoid-shaped hysteresis. Mechanical oscillations of the stereocilial bundle on the isolated hair cells also show hysteresis (Martin and Hudspeth, 1999). The force and work measured from the hair-bundle motion show loops when evaluated as a function of the displacement. These evidences suggest that hysteresis may be essential to the mechanical sensory processes in the inner ear where hair cell transduction is the fundamental element.

Negative damping has been proposed to be the active component in cochlear mechanisms (Neely and Kim, 1983), but has never been measured experimentally. A van der Pol limit-cycle oscillator model, suggested for the generation of spontaneous OAE (Long *et al.*, 1991), is essentially a dynamical system with negative damping similar to the hysteresis model described above. This implies that a common mechanism is responsible, at least in part, for the generation of different OAEs. The hysteresis and negative damping in cochlear mechanics may arise from the active motor elements within the stereocilia and lateral wall of the OHCs. Adaptation of myosin motors in the hair bundle can shift the f_{Tr} in the direction of the displacement (Eatock, 2000), thus indicating hysteresis. There are two major adaptive mechanisms: fast and slow. Slow adaptation may relate to the automatic gain control of the transducer to accommodate the wide range of input sound level as shown in the present study. Fast adaptation and resulting hysteresis could be responsible for the cycle-by-cycle amplification in the hair cell response. A recent study (Kennedy *et al.*, 2003) shows that fast adaptation presents in the mammalian OHC transduction mediated by the stereocilia. The observed double MP may reflect this type of adaptation for a given primary level. In addition to the compressive nonlinearity imposed by the sigmoid-shaped f_{Tr}, hysteresis could be another source for the generation of DPs.

Prestin, the OHC membrane motor, shows hysteresis or memory (Santos-Sacchi *et al.*, 2001a; Santos-Sacchi and Navarrete, 2002) if biased with a voltage pulse. Similar to the separation of the MPs of the CDT, these authors demonstrated a peak-shift of the cell membrane capacitance that governs the voltage-dependent force production. The electromotility of the OHC mediated by prestin is also tension- or stress-dependent (Santos-Sacchi *et al.*, 2001b). Effect of the

low-frequency bias tone is equivalent to alterations in the tension or stress on the cochlear transducer, and thus in turn modulates the motile response of the OHCs. The absolute value of the C factor derived from the model shows similar peak-shift [Fig. 10(d)] compared to the nonlinear capacitance of the isolated OHCs while manipulating the loads on the cells (Iwasa, 1993; Gale and Ashmore, 1994; Kakehata and Santos-Sacchi, 1995). The charge-movement or electrical current across the OHC membrane and the induced mechanical response of the cell demonstrate properties of piezoelectric materials (Weitzel *et al.*, 2003), where hysteresis is a typical nonlinear characteristic (Weissmüller *et al.*, 2003). If the active F_d is determined by the intrinsic electromechanical properties of the transducer, then it is advantageous for the OHCs to amplify high-frequency sound, since the velocity of vibration is greater. Indeed, the negative-damping model showed greater H for higher primary frequencies.

VI. SUMMARY AND CONCLUSION

A time-domain method has been developed to obtain a cochlear f_{Tr} using low-frequency modulated DPOAEs. The results of the experiment showed that the CDT magnitude as a function of bias level demonstrated a double MP consisting of two typical shapes of the $|f_{Tr}^{(3)}|$. The center peaks of the CDT double MP occurred at more positive or negative bias pressures when the bias pressure was rising or falling, respectively. This double MP suggested that the cochlear f_{Tr} had hysteresis and produced phase-related changes in the output that depend on the loading or unloading of the transducer. An important feature of the cochlear hysteresis was the counterclockwise traversal, which was opposite to most mechanical systems. To explore the underlying mechanism of the double MP and hysteresis, a model based on the second-order f_B with negative damping as feedback component was proposed. This model was able to produce the CDT double MP and counterclockwise hysteresis. Moreover, the CDT double MPs and the associated hysteresis were characterized with a set of physiologic indices. The model parameters were also obtained from fitting the experimental data. These physiologic indices and model parameters showed steeper f_{Tr} slope and higher feedback gains at lower signal levels. Observations from the present study suggest that cochlear f_{Tr} is dynamically nonlinear, hysteretic, and active. Both indices of the CDT double MP and model parameters could be used to access cochlear f_{Tr} and its hysteresis. Low-frequency biasing provides a nondestructive method to evaluate the mechanical properties of a remote system. Therefore, low-frequency modulation of DPOAEs could make *in vivo* evaluation of cochlear dynamics in humans possible.

ACKNOWLEDGMENTS

We thank David Mills, Wen Liu, and David Lerner for their suggestions on the experiments and valuable comments on an earlier version of the manuscript. This study was supported by the National Institute on Deafness and Other Communication Disorders of NIH, Grant Nos. R03 DC006165 and R01 DC02117.

Avan, P., Bonfils, P., Gilain, L., and Mom, T. (**2003**). ''Physiopathological significance of distortion-product otoacoustic emissions at $2f_1-f_2$ produced by high- versus low-level stimuli,'' J. Acoust. Soc. Am. **113**, 430–441.

Bendat, J. S., Palo, P. A., and Coppolino, R. N. (**1992**). ''A general identification technique for nonlinear differential equations of motion,'' Probab. Eng. Mech. **7**, 43–61.

Bian, L., and Chertoff, M. E. (**1998**). ''Differentiation of cochlear pathophysiology in ears damaged by salicylate or a pure tone using a nonlinear systems identification technique,'' J. Acoust. Soc. Am. **104**, 2261–2271.

Bian, L., and Chertoff, M. E. (**2001**). ''Distinguishing cochlear pathophysiology in 4-aminopyridine and furosemide treated ears using a nonlinear systems identification technique,'' J. Acoust. Soc. Am. **109**, 671–685.

Bian, L., Chertoff, M. E., and Miller, E. (**2002**). ''Deriving a cochlear transducer function from low-frequency modulation of distortion product otoacoustic emissions,'' J. Acoust. Soc. Am. **112**, 198–210.

Brownell, W. E., Bader, C. R., Bertrand, D., and de Ribaupierre, Y. (**1985**). ''Evoked mechanical response of isolated cochlear outer hair cells,'' Science **227**, 194–196.

Chertoff, M. E., Miller, E., and Bian, L. (**2001**). ''Properties and quantification of linear and nonlinear systems,'' Semin. Hear. **22**, 325–338.

Choi, C.-H., Chertoff, M. E., and Yi, X. (**2002**). ''Characterizing cochlear mechano-electric transduction with a nonlinear system identification technique: The influence of the middle ear,'' J. Acoust. Soc. Am. **112**, 2898–2909.

Cody, A. R., and Russell, I. J. (**1992**). ''Effects of intense acoustic stimulation on the nonlinear properties of mammalian hair cells,'' in *Noise-Induced Hearing Loss*, edited by A. L. Dancer, D. Henderson, R. J. Salvi, and R. P. Hamernik (Mosby Year Book, St. Louis, MO), pp. 11–27.

Cooper, N. P. (**1996**). ''Two-tone suppression in cochlear mechanics,'' J. Acoust. Soc. Am. **99**, 3087–3098.

Davis, H. (**1983**). ''An active process in cochlear mechanics,'' Hear. Res. **9**, 79–90.

Dirckx, J. J. J., Decraemer, W. F., von Unge, M., and Larsson, Ch. (**1998**). ''Volume displacement of the gerbil eardrum pars flaccida as a function of middle ear pressure,'' Hear. Res. **118**, 35–46.

Dorn, P. A., Konrad-Martin, D., Neely, S. T., Keefe, D. H., Cyr, E., and Gorga, M. P. (**2001**). ''Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears,'' J. Acoust. Soc. Am. **110**, 3119–3131.

Eatock, R. A. (**2000**). ''Adaptation in hair cells,'' Annu. Rev. Neurosci. **23**, 285–314.

Evans, B. N., Hallworth, R., and Dallos, P. (**1991**). ''Outer hair cell electromotility: The sensitivity and vulnerability of the dc component,'' Hear. Res. **52**, 288–304.

Fahey, P. F., Stagner, B. B., Lonsbury-Martin, B. L., and Martin, G. K. (**2000**). ''Nonlinear interaction that could explain distortion product interference response areas,'' J. Acoust. Soc. Am. **108**, 1786–1802.

Faulstich, M., and Kössl, M. (**2000**). ''Evidence for multiple DPOAE components based upon group delay of the $2f_1-f_2$ distortion in the gerbil,'' Hear. Res. **140**, 99–110.

Frank, G., and Kössl, M. (**1997**). ''Acoustic and electrical biasing of the cochlear partition. Effects on the acoustic two tone distortions f_2-f_1 and $2f_1-f_2$,'' Hear. Res. **113**, 57–68.

Gale, J. E., and Ashmore, J. F. (**1994**). ''Charge displacement induced by rapid stretch in the basolateral membrane of the guinea-pig outer hair cell,'' Proc. R. Soc. London, Ser. B **255**, 243–249.

Gale, J. E., and Ashmore, J. F. (**1997**). ''An intrinsic frequency limit to the cochlear amplifier,'' Nature (London) **389**, 63–66.

Geisler, C. D., and Nuttall, A. L. (**1997**). ''Two-tone suppression of basilar membrane vibrations in the base of the guinea pig cochlea using 'low-side' suppressors,'' J. Acoust. Soc. Am. **102**, 430–440.

Gold, T. (**1948**). ''Hearing. II. The physical basis of the action of the cochlea,'' Proc. R. Soc. London, Ser. B **135**, 492–498.

Hao, L. F., and Khanna, S. M. (**2000**). ''Mechanical nonlinearity in the apical turn of the guinea pig organ of Corti,'' Hear. Res. **148**, 31–46.

He, D. Z. Z., and Dallos, P. (**1999**). ''Somatic stiffness of cochlear outer hair cells is voltage-dependent,'' Proc. Natl. Acad. Sci. U.S.A. **96**, 8223–8228.

Hudspeth, A. J. (**1989**). ''How the ear's works work,'' Nature (London) **341**, 397–404.

Iwasa, K. H. (**1993**). ''Effect of stress on the membrane capacitance of the auditory outer hair cell,'' Biophys. J. **65**, 492–498.

Kakehata, S., and Santos-Sacchi, J. (**1995**). ''Membrane tension directly

shifts voltage dependence of outer hair cell motility and associated gating charge,'' Biophys. J. **68**, 2190–2197.
Kemp, D. T. (**1978**). ''Stimulated acoustic emissions from within the human auditory system,'' J. Acoust. Soc. Am. **64**, 1386–1391.
Kemp, D. T. (**1979**). ''Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea,'' Arch. Oto-Rhino-Laryngol. **224**, 37–45.
Kennedy, H. J., Evans, M. G., Crawford, A. C., and Fettiplace, R. (**2003**). ''Fast adaptation of mechanoelectrical transducer channels in mammalian cochlear hair cells,'' Nat. Neurosci. **6**, 832–836.
Kirk, D. L. (**2002**). ''Interaction between adenosine triphosphate and mechanically induced modulation of electrically evoked otoacoustic emissions,'' J. Acoust. Soc. Am. **111**, 2749–2758.
Kirk, D. L., and Yates, G. K. (**1998**). ''Enhancement of electrically evoked otoacoustic emissions associated with low-frequency stimulus bias of the basilar membrane towards scala vestibuli,'' J. Acoust. Soc. Am. **104**, 1544–1554.
Kössl, M., and Russell, I. J. (**1992**). ''The phase and magnitude of hair cell receptor potentials and frequency tuning in the guinea pig cochlea,'' J. Neurosci. **12**, 1575–1586.
Kros, C. J., Rüsch, A. R., and Richardson, G. P. (**1992**). ''Mechano-electrical transducer currents in hair cells of the cultured neonatal mouse cochlea,'' Proc. R. Soc. London, Ser. B **421**, 185–193.
Lakes, R. S. (**1999**). *Viscoelastic Solids* (CRC, New York), pp. 63–109.
Lee, C.-Y., and Rosowski, J. J. (**2001**). ''Effects of middle-ear static pressure on pars tensa and pars flaccida of gerbil ears,'' Hear. Res. **153**, 146–163.
Liberman, M. C., Gao, J., He, D. Z. Z., Wu, X., Jia, S., and Zuo, J. (**2002**). ''Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier,'' Nature (London) **419**, 300–304.
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.
Lukashkin, A. N., and Russell, I. J. (**1999**). ''Analysis of the f_2-f_1 and $2f_1-f_2$ distortion components generated by the hair cell mechanoelectrical transducer: Dependence on the amplitudes of the primaries and feedback gain,'' J. Acoust. Soc. Am. **106**, 2661–2668.
Lukashkin, A. N., Lukashkina, V. A., and Russell, I. J. (**2002**). ''One source for distortion product otoacoustic emissions generated by low- and high-level primaries,'' J. Acoust. Soc. Am. **111**, 2740–2748.
Martin, G. K., Villasuso, E. I., Stagner, B. B., and Lonsbury-Martin, B. L. (**2003**). ''Suppression and enhancement of distortion-product otoacoustic emissions by interference tones above f_2. II. Findings in humans,'' Hear. Res. **177**, 111–122.
Martin, G. K., Lonsbury-Martin, B. L., Probst, R., Scheinin, S. A., and Coats, A. C. (**1987**). ''Acoustic distortion products in rabbit ear canal. II. Sites of origin revealed by suppression contour and pure-tone exposures,'' Hear. Res. **28**, 191–208.
Martin, P., and Hudspeth, A. J. (**1999**). ''Active hair-bundle movements can amplify a hair cell's response to oscillatory mechanical stimuli,'' Proc. Natl. Acad. Sci. USA **96**, 14306–14311.
Mills, D. M. (**2000**). ''Frequency responses of two- and three-tone distortion product otoacoustic emissions in Mongolian gerbil,'' J. Acoust. Soc. Am. **107**, 2586–2602.
Mills, D. M. (**2002**). ''Interpretation of standard distortion product otoacoustic emission measurements in light of the complete parametric response,'' J. Acoust. Soc. Am. **112**, 1545–1560.
Mills, D. M., and Rubel, E. W. (**1997**). ''Development of distortion product emissions in the gerbil: 'filter' response and signal delay,'' J. Acoust. Soc. Am. **101**, 395–411.
Nakajima, H. H., Hubbard, A. E., and Mountain, D. C. (**2000**). ''Effects of acoustic trauma on acoustic enhancement of electrically evoked otoacoustic emissions,'' J. Acoust. Soc. Am. **107**, 2603–2614.
Neely, S. T., and Kim, D. O. (**1983**). ''An active cochlear model showing sharp tuning and high sensitivity,'' Hear. Res. **9**, 123–130.
Overstreet III, E. H., and Ruggero, M. A. (**2002**). ''Development of wide-band middle ear transmission in the Mongolian gerbil,'' J. Acoust. Soc. Am. **111**, 261–270.
Parthasarathi, A. A., Groth, K., Zheng, J., and Nuttall, A. L. (**2003**). ''Effect of current stimulus on *in vivo* cochlear mechanics,'' J. Acoust. Soc. Am. **113**, 442–452.
Patuzzi, R. (**1996**). ''Cochlear micromechanics and macromechanics,'' in *The Cochlea*, edited by P. Dallos, A. N. Popper, and R. R. Fay (Springer, New York), pp. 186–257.
Patuzzi, R., and Moleirinho, A. (**1998**). ''Automatic monitoring of mechanoelectrical transduction in the guinea pig cochlea,'' Hear. Res. **125**, 1–16.
Peters, R. D., and Pritchett, T. (**1997**). ''The not-so-simple harmonic oscillator,'' Am. J. Phys. **65**, 1067–1073.
Ren, T., and Nuttall, A. L. (**1998**). ''Acoustical modulation of electrically evoked otoacoustic emission in intact gerbil cochlea,'' Hear. Res. **120**, 7–16.
Robles, L., and Ruggero, M. A. (**2001**). ''Mechanics of the mammalian cochlea,'' Physiol. Rev. **81**, 1305–1352.
Santos-Sacchi, J., and Navarrete, E. (**2002**). ''Voltage-dependent changes in specific membrane capacitance caused by prestin, the outer hair cell lateral membrane motor,'' Pfluegers Arch. **444**, 99–106.
Santos-Sacchi, J., Shen, W., Zheng, J., and Dallos, P. (**2001a**). ''The outer hair cell lateral membrane motor, prestin, shows hysteresis,'' Biophys. J. **80**, 346a.
Santos-Sacchi, J., Shen, W., Zheng, J., and Dallos, P. (**2001b**). ''Effects of membrane potential and tension on prestin, the outer hair cell lateral membrane motor protein,'' J. Physiol. (London) **531**, 661–666.
Schmiedt, R. A., and Zwislocki, J. J. (**1977**). ''Comparison of sound-transmission and cochlear-microphonic characteristics in Mongolian gerbil and guinea pig,'' J. Acoust. Soc. Am. **61**, 133–149.
Scholz, G., Hirschfelder, A., Marquardt, T., Hensel, J., and Mrowinski, D. (**1999**). ''Low-frequency modulation of the $2f_1-f_2$ distortion product otoacoustic emissions in the human ears,'' Hear. Res. **130**, 189–196.
Shera, C. A., and Guinan, J. J. (**1999**). ''Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs,'' J. Acoust. Soc. Am. **105**, 782–798.
Spoendlin, H. (**1985**). ''Anatomy of cochlear innervation,'' Am. J. Otolaryngol. **6**, 453–467.
Tongue, B. H. (**2002**). *Principles of Vibration*, 2nd ed. (Oxford U.P., New York), pp. 69–151.
Weissmüller, J., Viswanath, R. N., Kramer, D., Zimmer, P., Würschum, R., and Gleiter, H. (**2003**). ''Charge-induced reversible strain in a metal,'' Science **300**, 312–315.
Weitzel, E. K., Tasker, R., and Brownell, W. E. (**2003**). ''Outer hair cell piezoelectricity: Frequency response enhancement and resonance behavior,'' J. Acoust. Soc. Am. **114**, 1462–1466.
Whitehead, M. L., Stagner, B. B., Martin, G. K., and Lonsbury-Martin, B. L. (**1996**). ''Visualization of the onset of distortion-product otoacoustic emissions, and measurement of their latency,'' J. Acoust. Soc. Am. **100**, 1663–1679.
Yates, G. K. (**1990**). ''Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions,'' Hear. Res. **50**, 145–162.
Zimmermann, U., and Fermin, C. (**1996**). ''Shape deformation of the organ of Corti associated with length changes of outer hair cell,'' Acta Otolaryngol. (Stockh.) **116**, 395–400.

Dispersive instability and its minimization in time-domain computation of steady-state responses of cochlear models

Jack Xin[a)]
Department of Mathematics and ICES, University of Texas at Austin, Austin, Texas 78712

(Received 2 September 2003; revised 3 February 2004; accepted 18 February 2004)

Dispersive instability appears in time-domain solutions of classical cochlear models. In this letter, a derivation of optimal initial data is presented to minimize the effect of instability. A second-order accurate implicit boundary integral method is introduced. Numerical solutions of two-dimensional models show that the optimal initial conditions work successfully in time-domain steady-state computations for both the zero Neumann and zero Dirichlet fluid pressure boundary conditions at the helicotrema. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1699393]

PACS numbers: 43.64.Kc [JBS] Pages: 2173–2177

I. INTRODUCTION

Time-domain computation of cochlear models is necessary for computing basilar-membrane (BM) responses in the presence of nonlinearities.[1–4] However, time-domain solutions are prone to dispersive instability[4] even at the linear level. In Sec. II, illustrative numerical examples are given on such an instability arising in time-domain solutions of classical linear cochlear models, where frequency-domain solutions are perfectly fine.[5]

It is useful to investigate how to do away with the instability and capture steady states dynamically as fast in time as possible. The objective here is to come up with a robust and effective approach without modifying models or specializing model parameters. Though the dispersiveness of model equations is the source of the problem, the cure lies in choosing a class of initial data so that the instability can be minimized and steady-state responses effectively computed.

In Sec. III, precise conditions of such optimal initial data are derived for the two-dimensional model under zero Neumann fluid pressure boundary condition at the helicotrema. In Sec. IV, an associated second-order accurate implicit boundary integral method is presented and analyzed. Numerical results show that the optimal initial data work effectively for both zero Neumann and zero Dirichlet pressure boundary conditions. Conclusions are given in Sec. V.

II. DISPERSIVE INSTABILITY

Classical one-dimensional transmission line model reads[6]

$$p_{xx} - N u_{tt} = 0, \quad x \in (0,L), \tag{2.1}$$

$$p = m u_{tt} + r u_t + s(x) u, \tag{2.2}$$

where p is the fluid pressure difference across the basilar membrane (BM), u the BM displacement, L the longitudinal length of BM with stapes located at $x=0$, and helicotrema at $x=L$; N a constant equal to fluid density times the ratio of BM width and scala cross section; m,r,s are the mass, damping, and stiffness of BM per unit area, respectively, with m and r being positive constants, and s a known function of x. The scala fluid motion obeys incompressible Stokes system with constant density.

Equations (2.1)–(2.2) are evolved in time with the boundary and initial data[2]

$$p_x(0,t) = p_{\text{in}}(t), \quad p(L,t) = 0, \tag{2.3}$$

$$u(x,0) = 0, \quad u_t(x,0) = 0, \tag{2.4}$$

where p_{in} is proportional to stapes acceleration driven by input signal.

The classical two-dimensional model replaces (2.1) by the Laplace equation on p thanks to the incompressibility of fluids[3,5]

$$p_{xx} + p_{zz} = 0, \quad (x,z) \in (0,L) \times (0,H), \tag{2.5}$$

subject to boundary conditions

$$p_x|_{x=0} = -2\rho\xi(t), \quad p|_{x=L} = 0, \tag{2.6}$$

$$p_z|_{z=0} = 2\rho u_{tt}, \quad p_z|_{z=H} = 0, \tag{2.7}$$

where H is the vertical height, ρ the fluid volume density, $\xi(t)$ the stapes acceleration. The initial condition (2.4) remains the same.[3]

The time-dependent systems above are known to be dispersive. If one ignores the boundary conditions at $x=0$, L, real dispersion relation is obtained in closed form when setting $r=0$, s to a positive constant, and looking for plane waves. A dispersive wave property is that spatial long waves move faster than short waves.[4]

For single-frequency input, $\sin(2\pi f t + \varphi)$, the time-dependent solutions lock onto the input frequency in time and eventually develop a spatial profile. It is generally expected that the steady state is a time harmonic solution, or the imaginary part of the complex solution $\mathbf{A}(x,z;f)\exp\{i(2\pi f t + \varphi)\}$, known as the frequency-domain solution.

However in numerical computations, time-domain solutions $u(x,t)$ contain additional dispersive waves which may outgrow the desired response.[4] As an illustration, in Fig. 1, we show the BM displacement $u(x, t = 20\ \text{ms})$, when the input is $\sin(2\pi f t)$, $f = 3$ kHz, computed by a second-order accurate implicit finite difference method[4] of the one-

[a)]Electronic mail: jxin@math.utexas.edu.

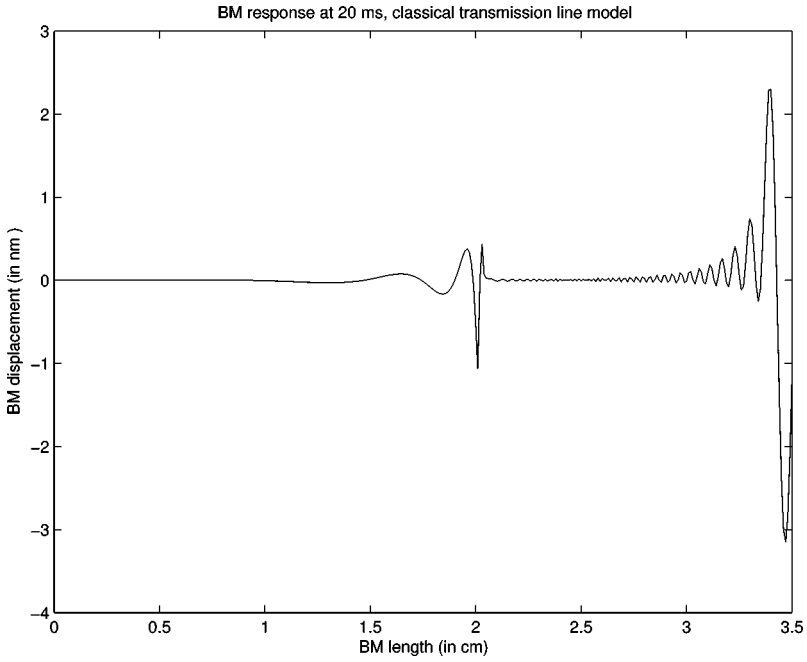


FIG. 1. BM displacement (in nm) at 20 ms with a sinusoidal input of 40 dB (SPL) and 3 kHz, computed by an implicit second-order finite difference method from the classical transmission line model (2.1)–(2.4). Mass density $m=0.05$ g/cm^2, damping constant $r=100$ dyn $\cdot$ s/cm^3, stiffness $s(x)$ as in (4.7). Midear factor is included. Steady state has formed to the left of the characteristic location (about 2 cm), yet to the right of it, dispersive instability is developed and persists afterward.

dimensional transmission line model with standard choice of coefficients. The dispersive waves are developed in the middle of the domain. They eventually travel to $x=L$ and remain there at later time. Figure 2 is a computed BM displacement at 10 ms from the two-dimensional model, using a second-order accurate implicit boundary integral method. The solution has reached steady state, yet the dispersive tail wave stays. The phenomenon is persistent under grid refinement, indicating that it is intrinsic to analytical solutions. Alternative numerical methods also reveal the same instability.

The key idea for developing a robust method to compute steady states in the time domain is to select suitable initial data different from the conventional data in (2.4) so as to minimize the growth of dispersive waves in the wake of steady states. Such optimal initial data shall depend on the input signal.

III. OPTIMAL INITIAL DATA

It is not obvious at all how to choose initial data in the systems (2.1)–(2.7) better than the natural all-zero data. Let us instead consider system (2.5)–(2.7), with the Dirichlet boundary condition $p|_{x=L}=0$ replaced by the zero Neumann boundary condition $p_x|_{x=L}=0$, or a rigid wall condition. At least for frequency-domain solutions, the resulting change is known to be insignificant; see the discussions in Neely[5] and references therein. However, there is a major difference in

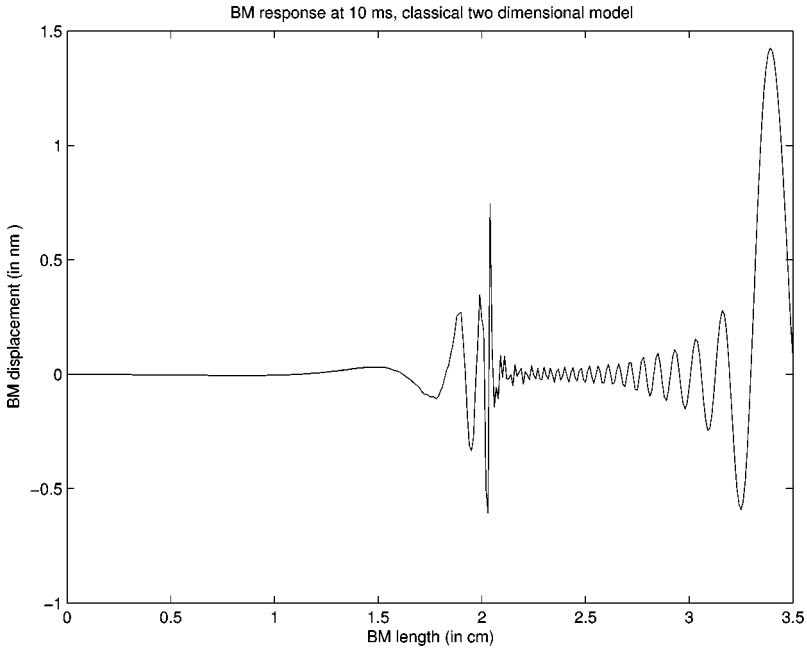


FIG. 2. BM displacement (in nm) at 10 ms with a sinusoidal input of 40 dB (SPL) and 3 kHz, computed by an implicit boundary integral method from the classical two-dimensional cochlear model (2.5)–(2.7). Model and numerical parameters are as in Sec. IV. Midear factor is included. Steady state has formed to the left of the characteristic location (about 2 cm), yet to the right of it, dispersive instability is developed and persists afterward.

mathematical formulation. A solvability condition must be imposed on the boundary data for the Neumann Laplacian, as a consequence of Green's theorem. It is

$$\int_0^L u_{tt}(x,t)dx = \xi(t)H, \tag{3.1}$$

which is preserved in time provided Eq. (2.2) is constrained as

$$p(x,0,t)+\lambda(t) = mu_{tt}+ru_t+s(x)u, \tag{3.2}$$

where $\lambda(t)$ is a time-dependent Lagrange multiplier so that (3.1) holds. It is clear that

$$\lambda(t) = -\frac{1}{L}\int_0^L [p(x,0,t)-ru_t-s(x)u]dx+\frac{m}{L}\xi(t)H. \tag{3.3}$$

The role of $\lambda(t)$ is to ensure that the evolution of u is consistent with the solvability condition (3.1) for all time. With this formulation, a time-domain method can be devised. Before we do that, let us extract some information about u from Eq. (3.1).

Integrating (3.1) in time twice, we find

$$\int_0^L u_t(x,t)dx = \alpha_1 + H\int_0^t \xi(\tau)d\tau, \tag{3.4}$$

and

$$\int_0^L u(x,t)dx = \alpha_2+\alpha_1 t + H\int_0^t dt'\int_0^{t'}\xi(\tau)d\tau, \tag{3.5}$$

where

$$\alpha_1 = \int_0^L u_t(x,0)dx, \tag{3.6}$$

$$\alpha_2 = \int_0^L u(x,0)dx. \tag{3.7}$$

For a general single-tone signal, $\xi(t)=A\sin(\omega t+\varphi)$, $\omega = 2\pi f$. The right-hand side of (3.5) is integrated exactly to give

$$\int_0^L u(x,t)dx = t(AH\omega^{-1}\cos\varphi+\alpha_1)+AH\omega^{-2} \times(\sin\varphi-\sin(\omega t+\varphi))+\alpha_2. \tag{3.8}$$

There is linear growth of $\int_0^L u(x,t)dx$ or instability unless

$$\alpha_1 = -AH\omega^{-1}\cos\varphi. \tag{3.9}$$

Interestingly, unless $\cos\varphi=0$, α_1 *should not equal zero*! This partially explains why instability is generic when computing the seemingly perfect system (2.1)–(2.7), assuming (and we shall see this is supported by numerics later) that the zero Neumann boundary condition at $x=L$ does little to the instability which arises in the interior of the computational domain.

Next, we choose α_2 as

$$\alpha_2 = -AH\omega^{-2}\sin\varphi, \tag{3.10}$$

so that

$$\int_0^L u(x,t)dx = -AH\omega^{-2}\sin(\omega t+\varphi), \tag{3.11}$$

has mean zero over $[0,T]$, where T equals the period f^{-1}.

Such a choice of α_2 also minimizes the L^2 norm square of $\int_0^L u(x,t)dx$ in the time interval $[0,T]$. In fact, it is easy to check that (3.11) is satisfied exactly by the time harmonic steady-state solution. For general signals, the optimality conditions on α_1 and α_2 can be derived by minimizing

$$\int_0^T\left[\alpha_2+\alpha_1 t+H\int_0^t dt'\int_0^{t'}\xi(\tau)d\tau\right]^2 dt,$$

over a large enough number T. Equivalently, one could do it frequency by frequency based on a spectral representation of the signal.

IV. NUMERICAL METHOD AND RESULTS

Let us consider the general single-tone input, and choose initial data as

$$u(x,0)=\alpha_2 g(x) \quad u_t = \alpha_1 g(x), \tag{4.1}$$

where $g(x)\in C^2([0,L])$, $g'(0)=g'(L)=0$, $\int_0^L g(x)dx=1$; and α_1, α_2 as in (3.9)–(3.10) so that (3.11) holds as a constraint for all time.

We shall compute with a boundary integral method by writing the pressure as a functional of u_{tt}, then reducing the problem to an integral-differential equation of u, and performing discretization. This is similar to the integral equation method in earlier works[1,3] except for the constraint (3.11). Inverting the Neumann Laplacian to solve for the pressure with the help of separation of variables, we obtain

$$p(x,0,t) = \frac{\rho\xi(t)}{L}((x-L^2)-H^2) -4\rho\sum_{k=1}^{\infty}\frac{\int_0^L u_{tt}\cos(k\pi x/L)dx}{k\pi\tanh(k\pi H/L)}\cos(k\pi x/L) \equiv p_0(x,t)+p_1(x,t) \tag{4.2}$$

The constrained u system is

$$p(x,0,t)+\lambda(t) = mu_{tt}+ru_t+s(x)u-\epsilon u_{xx}, \tag{4.3}$$

$$\int_0^L u(x,t)dx = -AH\omega^{-2}\sin(\omega t+\varphi), \tag{4.4}$$

where we have included a regularizing second-order stiffness term ϵu_{xx}, ϵ being a small, positive constant. The role of this second-order term is to allow us to impose zero Neumann boundary condition on u: $u_x(0,t)=u_x(L,t)=0$, which ensures that $u_{ttx}|_{x=0,L}=0$ so that the series expansion in (4.2) converges uniformly in $x\in[0,L]$. Mechanically, this term adds the longitudinal elastic tension on the membrane,[7] and was considered in BM models[8] earlier.

Now, we truncate the series in (4.2) with its first K terms, and discretize the integral in p_1 by composite trapezoidal rule. The semidiscrete system becomes

$$((mId+4\rho M)(u_{tt})^n)_j-\lambda^n = [p_0-ru_t-su+\epsilon u_{xx}]_j^n, \tag{4.5}$$

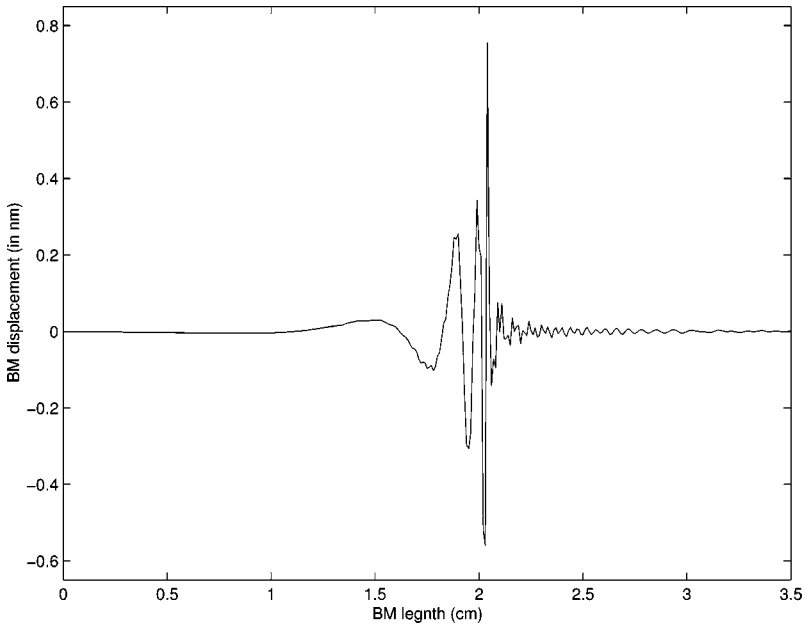


FIG. 3. BM displacement (in nm) at 10 ms with a sinusoidal input of 40 dB (SPL) and 3 kHz, computed by the implicit boundary integral method of the two-dimensional cochlear model with optimal initial data and zero Neumann data on pressure at $x=3.5$ cm. Midear factor is included. Steady state has formed to the left of the characteristic location (about 2 cm), and to the right of it, dispersive instability has been minimized.

where (j,n) denotes the grid points $(x,t)=(j(dx),n(dt))$, dx and dt being spatial and temporal grid steps; Id and M are $J\times J$ matrices; Id the identity, J total number of spatial grid points; $M=(m_{lj})$

$$m_{lj}=hc_j\sum_{k=1}^{K}\frac{\coth(k\pi H/L)}{k\pi}\cdot\cos(k\pi x_l/L)\cdot\cos(k\pi x_j/L),$$

j,l, range between 1 and J; $c_1=c_J=1/2$, $c_j=1$ if $1<j<J$.

Let $Q=\mathrm{diag}(1/2,1,...,1,1/2)$; then, the matrix QM is self-adjoint and non-negative. This is actually a property of the continuum operator from u_{tt} to p_2, which is compact on $L^2([0,L])$ and permits convergent finite dimensional approximations. The need for the diagonal matrix multiplier is due to trapezoidal rule. Hence, the operator $mId+4\rho M$ is invertible as a bounded (spatial) operator uniformly in $K\to\infty$.

The constrained system is further discretized as $(u_{tt})_j^n$ and $(u_t)_j^n$ are replaced by standard second-order central differencing, and

$$(u_{xx})_j^n=\tfrac{1}{4}\delta_x^2(u_j^{n+1}+2u_j^n+u_j^{n-1}),$$

where δ_x^2 is the standard spatial second-order differencing. Such a discretization leads to a stable implicit method so that time step dt is less restricted by dx and steady states are reached in shorter time evolution[4] than by explicit methods. Discretizing the integral in (4.4) by trapezoidal rule, the fully discrete system takes the block form

$$\begin{bmatrix} mId+4\rho M_1 & b_1 \\ b_2^T & 0 \end{bmatrix}\begin{bmatrix} u^{n+1} \\ \lambda^{n+1} \end{bmatrix}=\mathbf{B}(u^n,u^{n-1}), \tag{4.6}$$

where M_1 is symmetrizable by left multiplier Q and QM_1 is positive definite; $b_1=(1,...,1)^T\in R^J$, $b_2=(1/2,1,...,1,1/2)^T$

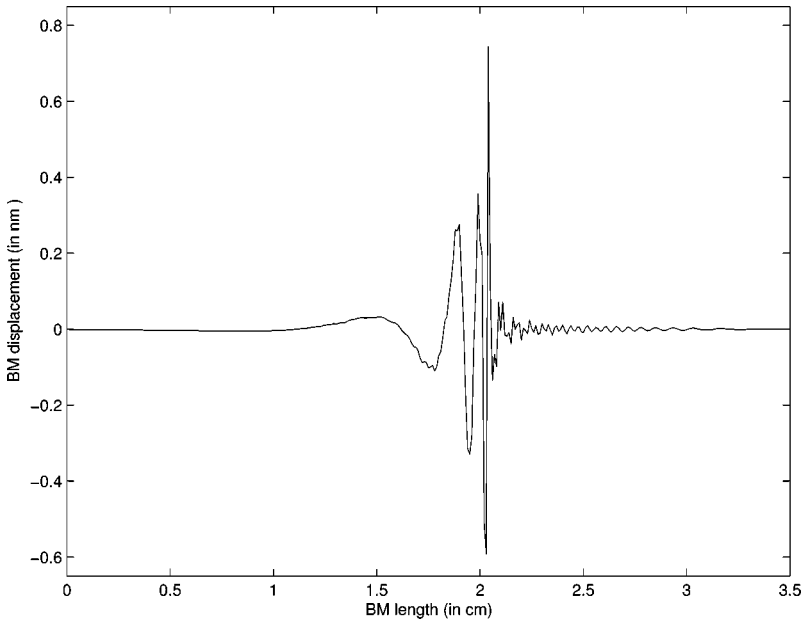


FIG. 4. BM displacement (in nm) at 10 ms with a sinusoidal input of 40 dB (SPL) and 3 kHz, computed by the implicit boundary integral method of the two-dimensional cochlear model with optimal initial data as in (5), and zero Dirichlet data on pressure at $x=3.5$ cm. Midear factor is included. Steady state has formed to the left of the characteristic location (about 2 cm), and to the right of it, dispersive instability has also been minimized by optimal initial data.

$\in R^J$, dots referring to 1's, T the transpose; and vector $\mathbf{B} \in R^{J+1}$ depends only on known values of u at previous time steps. The matrix M_1 is related to M as

$$M_1 = M + (4\rho)^{-1}\left(\frac{dt}{2}\cdot r\cdot Id - \frac{(dt)^2}{4(dx)^2}\cdot \epsilon\cdot M_2\right),$$

and $M_2=(m'_{i,j})$ is the tridiagonal matrix with -2 on the diagonal, and 1 on the off-diagonals, except $m'_{1,2}=m'_{J,J-1}=2$. The vector function $\mathbf{B}$ is:

$$\mathbf{B} = (mId + 4\rho M)(2u^n - u^{n-1}) + \frac{dt}{2} r u^{n-1}$$

$$+ (dt)^2 \frac{\epsilon}{4}\delta_x^2(2u^n + u^{n-1}) - (dt)^2 S u^n + (dt)^2 \mathbf{p}_0^n,$$

where S is the diagonal matrix with $s(j(dx))$'s on the diagonal, $j=1,2,...,J$; $\mathbf{p}_0^n$ has components $p_0(j(dx),n(dt))$, $j=1,2,...,J$. Also, zero Neumann boundary conditions are used when δ_x^2 involves the boundary points.

The block matrix in (4.6) is invertible. In fact, we have the identity

$$\begin{bmatrix} Q & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} mId + 4\rho M_1 & b_1 \\ b_2^T & 0 \end{bmatrix} = \begin{bmatrix} mQ + 4\rho Q M_1 & Qb_1 \\ b_2^T & 0 \end{bmatrix},$$

resulting in a symmetric block matrix of the same form upon noticing that $b_2 = Qb_1$, and Q, QM_1 are symmetric. As b_2 is nonzero and $mQ + 4\rho Q M_1$ is positive definite, the symmetric block matrix is invertible, and so is the block matrix in (4.6).

The method is well-defined with a standard initialization step. For the numerical computation reported here, the initial g function is: $g(x) = 60\max[x^2(1-x)^3, 0]$. The $s(x)$ is a functional fit of the Liberman data[9]

$$s(x) = 4\pi^2 m(0.456\exp(4.83(1-x/3.5)) - 0.45)^2, \tag{4.7}$$

in dyn/cm^3, the mass density $m=0.1$ g/cm^2, the damping constant $r=200$ dyn$\cdot$s/cm^3; the fluid density $\rho=1$ g/cm^3, $L=3.5$ cm, $H=0.1$ cm consistent with parameters in Neely.[5]

The spatial step $dx=0.01$ cm, time step $dt=0.001$ ms, truncation order $K=100$, $\epsilon=10^{-6}$ g/s^2. Smaller ϵ makes no noticeable difference in solutions. The input is a sinusoidal 3-kHz tone of 40 dB SPL, with phase shift $\varphi=0$. At 3 kHz, the midear filter provides a gain factor of about 33, which has been included in the solutions.

Figure 3 shows the computed BM displacement (in nm) at 10 ms with a sinusoidal input of 40 dB (SPL) and 3 kHz. Steady state has formed to the left of the characteristic location (about 2 cm), and to the right of it, dispersive instability has been minimized as our theory predicted. Refined computations and variation of input and model parameters indicate that the optimal initial data work in a robust fashion.

What is striking is that the optimal initial data also minimize the dispersive instability in the case of zero Dirichlet pressure boundary data at $x=3.5$ cm. This is demonstrated in Fig. 4, computed with a similar second-order implicit boundary integral method. Comparing with Fig. 2, we see that the solution captures well the steady-state response. Comparing with Fig. 3, we observe that there is no significant difference between imposing zero Dirichlet or zero Neumann pressure boundary data at $x=3.5$ cm, as in case of the frequency-domain solutions.[5] For this reason, the optimal initial data obtained theoretically from the zero Neumann case apply to the classical model (2.4)–(2.7) as well, though it is not yet clear how to find the exact optimal conditions in the Dirichlet case.

V. CONCLUSIONS

Time-domain computations of steady-state cochlear responses in classical models are subject to dispersive instability. Optimal conditions on initial data are derived and shown to effectively minimize the instability effect for both the zero Neumann and zero Dirichlet pressure boundary conditions at the helicotrema in two-dimensional models. In future work, the analytical conditions on the optimal initial data and the implicit boundary integral method may be extended to time-domain solutions of nonlinear and active cochlear models.[10,11]

ACKNOWLEDGMENTS

This work was supported in part by NSF Grant ITR-0219004, the Faculty Research Assignment Award at the University of Texas at Austin, and a Fellowship from the John Simon Guggenheim Memorial Foundation. I thank M. D. Lamar for preliminary computations, and Y-Y Qi for helpful remarks.

[1] J. B. Allen and M. M. Sondhi, "Cochlear macromechanics: Time domain solutions," J. Acoust. Soc. Am. **66**(1), 123–132 (1979).
[2] R. J. Diependaal, H. Duifhuis, H. W. Hoogstraten, and M. A. Viergever, "Numerical methods for solving one-dimensional cochlear models in the time domain," J. Acoust. Soc. Am. **82**(5), 1655–1666 (1987).
[3] R. J. Diependaal and M. A. Viergever, "Nonlinear and active two-dimensional cochlear models: Time-domain solution," J. Acoust. Soc. Am. **85**(2), 803–812 (1989).
[4] J. Xin, Y-Y. Qi, and L. Deng, "Time domain computation of a nonlinear nonlocal cochlear model with applications to multitone interaction in hearing," Comm. Math Sci. **1**(2), 211–227 (2003). Access on-line at: ⟨www.intlpress.com/CMS/journal⟩
[5] S. T. Neely, "Finite difference solution of a two-dimensional mathematical model of the cochlea," J. Acoust. Soc. Am. **69**(5), 1386–1393 (1981).
[6] M. M. Sondhi, *The Acoustical Inverse Problem for the Cochlea*, Lecture Notes in Biomath Vol. 43, edited by M. Holmes and L. Rubenfeld (Springer, Berlin, 1980), pp. 95–104.
[7] T. Rossing and N. Fletcher, *Vibration and Sound* (Springer, Berlin, 1995).
[8] L. Deng and C. D. Geisler, "Responses of auditory-nerve fibers to multiple-tone complexes," J. Acoust. Soc. Am. **82**(6), 1989–2000 (1987).
[9] M. C. Liberman, "The cochlear frequency map for the cat: Labeling auditory nerve fibers of known characteristic frequency," J. Acoust. Soc. Am. **72**, 1441–1449 (1982).
[10] C. Steele, G. Baker, J. Tolomeo, and D. Zetes, "Cochlear Mechanics," in *Biomedical Engineering Handbook*, edited by J. Bronzino (CRC Press, Boca Raton, FL, 1995), pp. 505–516.
[11] C. Geisler and C. Sang, "A cochlear model using feed-forward outer-hair-cell forces," Hear. Res. **86**, 132–146 (1995).

High-frequency electromotile responses in the cochlea

Karl Grosh
Department of Mechanical Engineering, The University of Michigan, 3124 G.G. Brown Building, Ann Arbor, Michigan 48109-2125

Jiefu Zheng and Yuan Zou
Oregon Hearing Research Center (NRC04), Department of Otolaryngology and Head & Neck Surgery, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239-3098

Egbert de Boer
Room D2-226, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands

Alfred L. Nuttall
Oregon Hearing Research Center (NRC04), Department of Otolaryngology and Head & Neck Surgery, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239-3098, and Kresge Hearing Research Institute, The University of Michigan, 1301 East Ann Street, Ann Arbor, Michigan 48109-0506

(Received 21 November 2003; revised 10 February 2004; accepted 16 February 2004)

Mammalian outer hair cells (OHCs) convert electrical energy into mechanical energy. The significance of this electromotility rests in the ability of the OHCs to modulate the vibrations of the cochlear partition *in vivo*. While high-frequency electromotility of isolated OHCs has been demonstrated at frequencies up to 100 kHz, a similar measure of the effect of OHC electromotility on motion of the sensory epithelium has not been made *in vivo*. In this study, *in vivo* electrical stimulation of the guinea pig cochlea is found to induce a mechanical response of the basilar membrane for frequencies to at least 100 kHz, nearly twice the upper limit of hearing for the guinea pig. The perfusion of salicylate in the cochlea reversibly reduces the electromotile response, indicating that an OHC-mediated process is the key contributor. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1695431]

PACS numbers: 43.64.Kc, 43.64.Me, 43.64.Jb [BLM] Pages: 2178–2184

I. INTRODUCTION

The outer hair cells (OHCs) of the mammalian cochlea are biologically unique because they are both sensors and actuators. In other words, OHCs function as mechano-electrical and electro-mechanical transducers. *In vitro* studies have shown that isolated OHCs are capable of mechanical oscillations at frequencies at least as high as 100 kHz in response to transmembrane sinusoidal electrical stimulation (Frank *et al.*, 1999). Electrical stimulation applied to the intracochlear fluids is known to elicit basilar-membrane motion (Xue *et al.*, 1995). Extra- and intracochlear electrical stimulation has been found to produce sound emissions measured in the external ear (Hubbard and Mountain, 1983; Nuttall, 1995; Xue, 1996) known as electrically evoked otoacoustic emissions (EEOAE). The upper-frequency limit of these emissions is about equal to the upper limit of hearing for that animal. In this paper, we investigate the frequency characteristics and test the upper-frequency limit of basilar-membrane (BM) velocity in response to sinusoidal intra- and extracochlear electrical stimulation *in vivo* and compare these results to the *in vitro* limits of isolated OHCs. We present results from control and postmortem experiments. These high-frequency results are used to determine mechanisms of electromotility in the cochlea. We discuss how these measurements could be used to estimate parameters in a mathematical model.

The precise micromechanical mechanisms for the *in vivo* OHC electromotility remain unknown, but their potential significance for hearing requires that they be able to modulate the vibrations of the cochlear partition *in vivo*. The present view is that the OHC mediates a cycle-by-cycle sound amplification in the following manner: (1) sound energy causes a traveling wave along the basilar membrane, displacing the components of the organ of Corti; (2) that this displacement leads to the deflection of OHC stereocilia, gating an ionic current that causes an oscillatory OHC membrane potential at the frequency of the stimulus; and (3) that OHCs convert the electrical energy into mechanical energy that feeds back with the correct phase to enhance vibration. This is the so-called active process of cochlear amplification (Dallos, 1992). The high-frequency motility of outer hair cells, encompassing (and surpassing) the highest physiologically relevant frequencies, provides for the frequency selectivity and sensitivity of mammalian hearing (see, e.g., Liberman *et al.*, 2002). While *in vitro* studies of high-frequency isolated OHC electromotility are available, there are no data on the high-frequency limit of electrical stimulation *in vivo*.

II. METHODS

A. Surgical preparation

We summarize the surgical preparation and animal handling used for this study, which are discussed in more detail

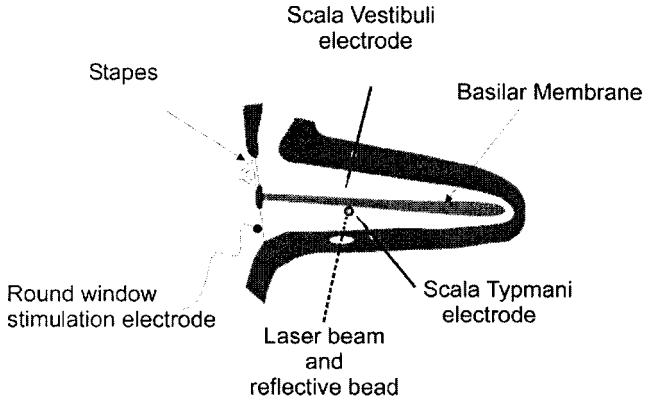


FIG. 1. Simplified (uncoiled) rendition of the cochlea showing the configuration for local or RW electrical stimulation. The local BM velocity was measured using a laser Doppler velocimeter. The scala tympani opening for the laser was also used to introduce the sodium salicylate.

in Parthasarathi *et al.*, 2003. Healthy young pigmented guinea pigs (250–400 g) were used in this study. The animals were anesthetized. A tracheotomy was performed after the animal's head was fixed to a headholder, and a ventilating tube inserted to ensure free breathing. The experimental protocols used were in accordance with the rules established by the Committee on Use and Care of Animals at the Oregon Health and Science University. After the bulla was opened to expose the cochlea, a silver wire (75-μm diameter) electrode was placed on the round window (RW).

A total of three openings was made in the cochlea [two in the *scala tympani* (ST) and one in the *scala vestibuli* (SV)]; see Fig. 1. An opening approximately 300 μm wide was made on the ST side of the cochlear basal turn to allow for measurement of BM velocity. The other two openings were approximately 75 μm in diameter and were made to insert the SV and ST electrodes into the perilymph.

The compound action potential (CAP) measured from the RW electrode, with reference to an Ag–AgCl ground electrode in the soft tissues of the neck, was used to obtain information on the N_1 detection threshold of the CAP at a given acoustic frequency. The threshold was used as an indicator of the cochlear sensitivity. For animals used in this study, the CAP loss at the first-turn best frequency (BF) was less than 20 dB. The high-frequency electromotility (above BF) seen in the present study was not critically impacted by a depressed CAP threshold. However, the response below and near to BF is reduced for animals whose CAP threshold was depressed after surgery.

B. Excitation and response measurement

Acoustic stimuli were delivered through a 1/2-in. Bruel & Kjaer condenser microphone coupled to the external ear through a speculum. Electrical stimuli were delivered to the cochlea from a custom-designed constant current unit (CCU). Electrical stimulation was delivered at two locations, at the RW and across the first cochlear turn. The wire electrode (Pt–Ir 75-μm diameter) cemented in position at the RW membrane was used to deliver the current. The return electrode for the RW electrode was a chlorided silver wire in the soft tissue of the neck. The electrical stimulation from the

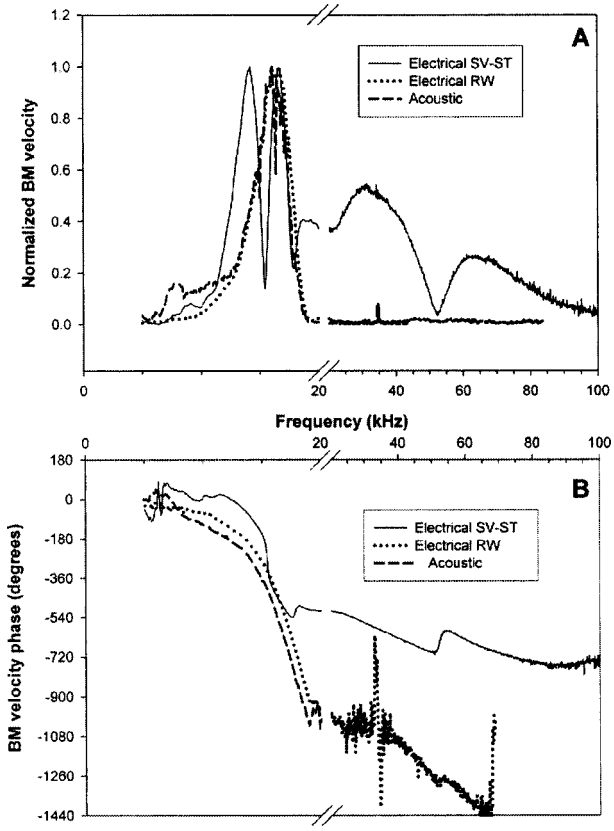


FIG. 2. Basilar-membrane velocity amplitude [panel (A)], normalized to maximum response in each case to acoustic, RW electrical, and local electrical (SV–ST) stimulation. In the case of electrical stimulation, the voltage sent to the constant current generator serves as the reference for the phase and the velocity was normalized to the applied current (GP 483). Phase relative to the voltage input to the excitation source is shown in panel (B). In response to bipolar electrical stimulation, the basilar-membrane response is seen to extend up to 100 kHz. An expanded scale of 5 kHz per tick mark is used for frequencies below 20 kHz.

RW–neck electrode pair is denoted as RW electrical stimulation in the sequel. Pt–Ir wires 50 μm in diameter were inserted into holes made in the bony cochlear wall of the first cochlear turn forming a bipolar pair across the cochlear duct from the SV to ST (see Fig. 1). We will denote this type of excitation as local bipolar stimulation. Voltage control to CCU and the speaker (via an amplifier) was from the oscillator output of a Stanford Research Systems (SR830) lock-in amplifier. BM velocity measurement was accomplished by directing the laser beam of a Polytec OFV 1102 through a compound microscope at a glass bead (20- or 3-μm diameter) placed onto the basilar membrane (see Fig. 1). The voltage output of the velocimeter was directed to the input of the lock-in amplifier.

III. RESULTS AND DISCUSSION

A. High-frequency basilar-membrane velocity response

Figure 2 shows the BM velocity response to three conditions (1) acoustic stimulation in the ear canal; (2) 35-μAmp rms electrical stimulation at the round window (RW); and (3) 100-μAmp rms local bipolar electrical stimulation. For acoustic stimulation (case 1), the mechanical response of

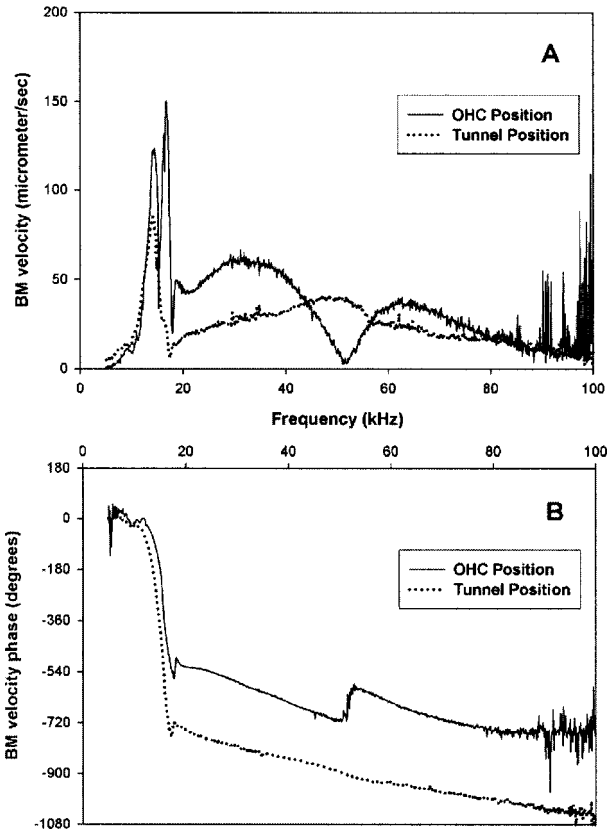


FIG. 3. Non-normalized basilar-membrane velocity amplitude [panel (A)] and phase [panel (B)] in response to local electrical stimulation (bipolar electrode placed across the SV/ST). These data are from the tunnel and OHC radial locations on the BM of GP 483. The BM velocity is seen to depend on the radial location of the bead.

the BM is tuned and has a peak response at about 17 kHz, the BF for this location. For frequencies above the BF the response falls to the noise floor [panel (A), dashed curve]. The corresponding phase curve [panel (B)] shows a pattern of increasing phase lag that is characteristic of a traveling wave. For electrical stimulation from the RW near the base of the cochlea (case 2), the pattern of response is similar to that from acoustic stimulation. This confirms earlier findings that electrical stimulation near the base of the cochlea results in forward-propagating traveling waves on the BM (Kirk and Yates, 1996; Nuttall *et al.*, 2001). In contrast, local bipolar electrical stimulation (case 3) produces a markedly different response (solid curves). This frequency response curve has both a more complex multipeaked pattern near the BF and the response extends to 100 kHz [panel (A)].

The high-frequency (above BF) response of the BM velocity was consistently measured in over 20 guinea pigs and showed remarkably little animal-to-animal variation. The amplitude response at this radial location (over the third row of OHCs) and longitudinal location shows a dip near 50 kHz and an approximate 180-deg phase reversal. This minimum in the response and the roughly 180-deg phase shift are indicative of a mechanical vibration mode, likely due to a resonance in the organ of Corti structures. The electrically evoked high-frequency response is present on all flexible portions in the radial direction of the BM. For instance, in Fig. 3, the BM velocity over the third row of OHCs is compared to the velocity measured over the tunnel region at the 17-kHz place (using two beads on the BM of the same animal). In both locations the high-frequency BM response to electrical stimulation is evident. The local electrical stimulation excites a radially varying response pattern that is most prominent above the BF of this location. The 50-kHz dip in the response seen over the OHCs is not present when measured at the tunnel region. Instead, a local maximum is seen at 50 kHz, and only a small phase shift is seen in these tunnel velocity data at 50 kHz. Due to the asymmetrical radial location of the OHC, actuation of the OHC will excite an asymmetrical structural mode. Since the asymmetrical BM mode has a higher resonance frequency than the more symmetric first mode, the excitation of this second structural mode will be more clearly seen at higher frequencies. For acoustic input, the asymmetric mode is not as prominent for two reasons. First, the force of the fluid pressure on the BM due to acoustic input has a spatial pattern that is more uniform in the radial direction than for excitation from OHCs, and therefore will more preferentially excite a symmetric BM mode. Second, since the acoustic pressure excites this region BM at frequencies at and below BF, the antisymmetric modal response is low as there is very little acoustical energy input to this mode near its resonance frequency (which is presumably above BF).

For frequencies less than the BF of the measurement location, bipolar electrical stimulation causes energy to be propagated basally to the stapes, setting up standing waves in the fluid between the excitation location and the base. The multiple peaks can be interpreted as resulting from interference among multiple traveling waves created by local forced stimulation of the BM by the OHCs. The smaller phase accumulation of the response evoked by local electrical stimulation [Fig. 2, panel (B)] is indicative of the presence forward and reverse wave propagation and inhomogeneous forcing (i.e., a combination of standing waves, traveling waves, and local forced responses) resulting in a lower phase shift than for a forward-traveling wave alone, such as is launched by acoustical and RW electrical excitation.

The constructive and destructive interference due to the wave propagation complicates our interpretation of the local mechanics at the measurement location. In a temporal bone preparation, Gummer *et al.* (1996) used electrical and acoustical excitation in much the same way as we have in the *in vivo* situation. There, they interpret the local amplitude minima and phase shifts around the BF as resonances in the TM and BM. This is one possible interpretation. There is no doubt that the local structural modal nature of the organ of Corti (OoC) and TM are involved. Over the third row of OHC, we see a beautiful antiresonance precisely between two peaks, one at the acoustical BF and another at a frequency nearly 1/2 octave below the BF (see Fig. 2). Some sort of dip at this frequency (between the BF and 1/2 octave below BF) is commonly seen in animals during electrical stimulation. A model that treats the BM, OoC, and TM structure as a locally reacting (i.e., no longitudinal stiffness) two degree of freedom oscillator, predicts that OHC forcing of the TM and BM, such as would occur from the electrical stimulation of the OHC, will cause a zero in the BM re-

sponse at the uncoupled resonance frequency of the TM (effective mass of TM combined with the effective stiffness of the TM and organ of Corti). If the cochlea were that simple, then the zero in the BM response would precisely locate the resonance of the TM. However, there is longitudinal stiffness in the cochlear structures, along with fluid–structure wave interference which, in our view, renders that interpretation ambiguous but worth further pursuit as a tool for identifying the *in vivo* properties of the OoC. While the dynamical system that is being excited by the bipolar stimulation is the same as that excited by the acoustical source, the location and type (electrical versus mechanical) are different. Hence, when estimating parameters of a model (e.g., Dimitriadis and Chadwick, 1999), these two experiments will provide independent data sets for such parameter estimation.

In a healthy cochlea, the BM response due to electrical or low-level acoustic input from the ear canal will be the sum fluid pressure plus OHC activity and other forces from the OoC unto the BM. We conjecture that at low input sound levels, the forcing from the OHCs would be more pronounced than for higher input acoustic excitation amplitudes, as they would account for a larger portion of the total response. One would then expect greatest asymmetry in the radial dependence of the BM velocity due to acoustic input to be seen in the most sensitive animals at low input levels, as is indicated in the results of Nilsen and Russell (2000). Electrical stimulation allows for the analysis of high-frequency spatial responses on the BM that are impossible with ear canal insonification.

It has been shown that the EEOAE response measured in the ear canal resulting from local, intracochlear electrical stimulation is bandlimited (Kirk and Yates, 1996; Nuttall *et al.*, 2001). These experiments show that the high-frequency cutoff of the EEOAE is correlated to the tonotopic location of the local electrical stimulation. For the guinea pig the frequency content of the EEOAE for RW electrical stimulation extends to 40 kHz, for first-turn excitation the limit is roughly 20 kHz, and for a turn 3 location the limit is 10 kHz (e.g., Nuttall *et al.*, 2001). This indicates that the electrical excitation from the bipolar electrodes is confined to a region close to the electrodes. If the electrical stimulation were to spread basally from the electrode location, evoked emissions would extend to higher frequencies by exciting the high-frequency, more basal OHCs. The EEOAE data from the literature are recounted because the velocity data due to local electrical stimulation in Fig. 2 show a local BM velocity response that is not bandlimited. Hence, the local mechanical response due to the putative OHC motor is active above BF, but those high-frequency waves are evanescent and do not reach the ear canal at measurable levels.

At ultrasonic frequencies above the BF, the response is dictated by local mechanical, electrical, and fluidic effects, rather than global wave propagation. The fluid loading at ultrasonic frequencies is mainly a local effect, since significant energy does not propagate away from the electrically excited OHC at ultrasonic frequencies. As such, this high-frequency excitation may provide a means of interrogating local health of a region of the cochlea as well as identifying parameters of a cochlear model that would be difficult to identify without high-frequency information. While the wave propagation effects in the cochlea are apparently simplified, there are other complications associated with using data from the high-frequency regime. The inertia of supporting cells becomes non-negligible and the structure of the organ of Corti will add additional loading. Electrically isolated OHCs will exhibit resonance-like behavior with resonance occurring at the 40–70-kHz region, as shown in the groundbreaking experiments of Frank *et al.* (1999) and in a recent model of OHC behavior (Weitzel *et al.*, 2003). The *in vivo* configuration is more complicated than the isolated hair cell configuration. The OHCs will, of course, become a component of the larger *in vivo* system which will possess its own dynamic characteristics. The details of some of the complicating factors (local mass, for example) are precisely what we may be interested in identifying.

B. Control experiments

We tested whether the measured velocity pattern is a result of a tracking problem between the BM vibration and the bead reflector (which is on the BM) (Nuttall and Dolan, 1996). The vibrations of 3- and 20-μm diameter beads were compared and no difference was found (3-μm bead data not shown). For one animal we successfully recorded BM vibration over the OHC without a reflective bead and also found a similar frequency response pattern, including the ''notch'' at 50 kHz (data not shown). The latter data support the earlier finding that beads can properly track BM motion (Cooper, 1999), although one other study concluded otherwise (Khanna *et al.*, 1998). The notch was a constant feature of the high-frequency portion of the spectrum for beads located over the rows of OHCs.

To demonstrate that the BM response is due to OHC electromotility, we applied sodium salicylate to the intracochlear fluids. Salicylate reduces OHC electromotility (Kakehata and Santos-Sacchi, 1996; Tunstall *et al.*, 1995). In Fig. 4, the perfusion of 20-μl of artificial perilymph containing 10-mM salicylate is shown to cause almost complete elimination of the electromotile response above 20 kHz. This effect was reversible. The BM velocity measurements shown in Fig. 4 are close to the tunnel region; hence, the response at 50 kHz has a peak consistent with Fig. 3. Salicylate reversibly reduces the high-frequency electromotility of the basilar membrane at concentrations that reversibly reduce isolated OHC motility. This implies that the OHCs are the main contributor to the high-frequency electromotility seen in this experiment.

While the data are consistent with OHC motility due to the presence of a basolateraltransmembrane protein, there may be a different mechanism at work. However, there is no experimental evidence to support significant electromotility in any of the other mammalian cochlear structures much above 1 kHz. Some limited electromotility at frequencies up to 20 kHz is evidenced in some cells, such as Chinese hamster ovary (CHO) cells (Ludwig *et al.*, 2001), although the force applied to an atomic force lever is 0.5 times that for prestin-transfected CHO cells, a factor of 10 less than for prestin-transfected human embryonic kidney cells, and there is a 180-deg phase difference between native and transfected

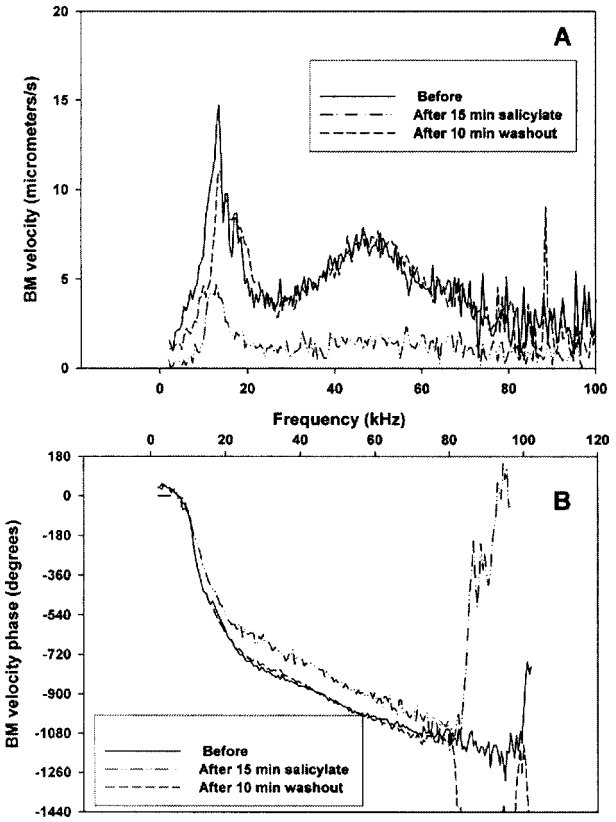


FIG. 4. BM velocity amplitude [panel (A)] and phase [panel (B)] before, during, and following washout of sodium salicylate from the cochlea (GP 1-06-03). Salicylate reversibly reduces the motility in response to electrical stimulation.

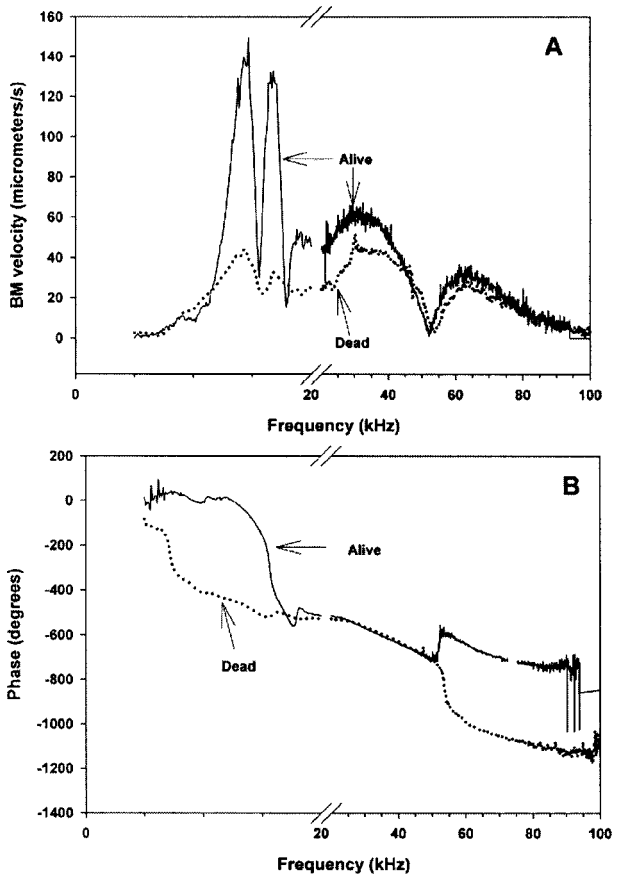


FIG. 5. Postmortem basilar-membrane velocity amplitude [panel (A)] and phase [panel (B)] compared to live data for the same animal (GP 483). Results are for local electrical bipolar stimulation. An expanded scale of 5 kHz per tick mark is used for frequencies below 20 kHz.

CHO cells. Another source of motility may come from the stereocilia. For instance, the hair bundles from the bullfrog sacculus evidence motility due to transepithelial electrical stimulation up to 100 Hz (Bozovic and Hudspeth, 2003). Calcium binding in a mechanically gated stereocilia channel has been postulated as a fast electromotile mechanism (possibly to the kHz range, e.g., Fettiplace *et al.*, 2001). Another possibility is that motility may arise from electrical stimulation of the lipid bilayer electric dipole in the cylindrical OHC (Petrov and Sachs, 2002), although this has not yet been confirmed in cochlear structures. OHC lipid motility (Petrov and Sachs, 2002) is partially blocked by salicylate. Note that spontaneous emissions at frequencies as high as 4 kHz from a gecko are suppressed by salicylate (Stewart and Hudpeth, 2000). Since there are no OHCs in the gecko, the implication is that salicylate also might inhibit hair-bundle motility. Therefore, we may be eliciting such a response from an unknown source in the cochlea via our electrical stimulations, effects that might be reduced by salicylate.

Electromotility of the cells is expected to be present postmortem as OHCs exhibit electromotility *in vitro*. Indeed, this is the case as shown in Fig. 5, where the BM velocity in response to bipolar stimulation is plotted before and immediately postmortem. For frequencies below the BF, the postmortem BM vibrations resemble the salicylate intoxicated frequency response. The reduction of BM velocity seen postmortem can be attributed to a disruption of the resting electrical and mechanical state of the cochlea by death. Because the endocochlear potential is nearly extinguished postmortem and hair cell active ionic pumps are nonfunctional, it is quite likely that the cell will be depolarized. Because of the change in polarization of the OHC, the mechanical state of the cell will be changed. The mechanical state characteristics include the OHC turgor pressure, the position of the stereocilia, and any tension/compression acting on the cell. Such an alteration of the transmembrane potential-to-mechanical length operating point of the OHCs could be postulated to either decrease or increase their overall effectiveness (or ''gain''). However, we argue that the *in vivo* resting configuration is somehow optimal for enhancing the response of the BM velocity in the normal physiological range (i.e., frequencies at and below BF). The response measured at this location for frequencies at and below BF involves, to some extent, all OHCs basal to this location. Hence, disrupting the *in vivo* operating condition is deleterious to global wave propagation and to the response near the BF, as seen in Fig. 5.

The *in vivo* and postmortem BM velocity response patterns above 40 kHz differ by only a small amount (Fig. 5). The notch at 50 kHz is retained in the postmortem velocity response. The smaller reduction is consistent with the view that very little energy propagates away from the locally excited OHCs at frequencies higher than BF in either the normal or postmortem case. Hence, this ultrasonic electromotility is less affected by alterations in the operating point of the

OHC as fewer OHCs are involved in this active process, which does not involve interaction with the traveling wave (and hence the basally located hair cells). The postmortem OHC evoked response depends on the slope of the motility versus voltage dependence. The postmortem OHC response might have shown an increase in electromotility (if moved to a more favorable point on the operating characteristic). However, we never saw an increase in the response postmortem.

IV. CONCLUSION

Our data show that sufficient force is produced *in vivo* to evoke vibrations at frequencies to 100 kHz and that a salicylate-sensitive process is involved. The dynamics of the fluid and mechanical environment surrounding the OHCs is likely modally rich at the ultrasonic frequencies measured in this study. Resonance structures, including a radially varying spatial dependence of the response and phase shift in the frequency response, are seen in the BM velocity. The ultrasonic response was found to show very little variability from animal to animal. Intracochlear bipolar electrical stimulation is a means of locally interrogating the *in vivo* OHC response, providing a complimentary alternative to *in vitro* preparations for determining electromotile processes and properties. In order to conclusively identify the physiological mechanism of electromotility in the cochlea, we expect that we will need to reconcile the results of *in vivo*, various types of *in vitro* experiments, and theoretical predictions.

Why perform high-frequency electrical stimulation? One goal is to identify the sources of electromotility in the cochlea. As of yet, only the somatic motility of the OHC has been demonstrated at these frequencies (Frank *et al.*, 1999). Other sources of electromotility are possible (such as the stereocilia), but in order for a convincing argument to be made for hair-bundle vibration as a main contributor, *in vitro* measurements on isolated cells showing high-frequency motility are needed. As mentioned by Fettiplace *et al.* (2001), there is no evidence *yet* of hair-bundle electromotility in mammalian OHCs. Once technical difficulties in such measurements are overcome, high-frequency bundle motility in mammalian OHCs might be seen. Arguing from the standpoint of conservation of genetic mechanisms, since activity is seen in lower vertebrates without OHCs, it might also be used by mammals in conjunction with somatic OHC motility. The argument for a combination of OHC somatic motility with some form of hair-bundle motility is still tenable. No matter the source of the electromotility, intracochlear electrical stimulation serves as a mechanically noninvasive way to excite the cochlea that is complimentary to acoustical excitation. The localized electrical stimulation is different spatially and temporally from acoustic stimulation directed through the ear canal. The frequencies of excitation are not limited by the best frequency at the measurement location, and the point of excitation is now localized along the BM. These high-frequency probes are used to examine high-frequency modal structures in the cochlea, which may be difficult to study using acoustical excitation (Nilsen and Russell, 2000) but might be important for cochlear sensitivity. Considering the response of the BM as a combination (albeit nonlinear) of forcing from the fluid pressure and active processes (e.g., from OHC forcing), bipolar electrical excitation is one means of teasing out the electrical and mechanical aspects of active processes. We note here that our bipolar stimulation is clearly different than electrical excitation that occurs normally in the cochlea. The intracochlear injection of current artificially fluctuates the local cochlear potentials. Finally, the local electrical excitation serves as an independent test of a cochlear model. For instance, one could fit a mathematical model to match standard acoustical excitation, then test the ability of the model to replicate electrically induced motion.

ACKNOWLEDGMENTS

The authors acknowledge the assistance of Dr. W. E. Brownell for his help in editing this manuscript. We also acknowledge the helpful comments of the reviewers. K.G. acknowledges the support of NIH-NIDCD RO1-DC04084 and ALN NIH-NIDCD RO1-DC00141.

Bozovic, D., and Hudspeth, A. J. (**2003**). ''Hair-bundle movements elicited by transepithelial electrical stimulation of hair cells in the sacculus of the bullfrog,'' Proc. Natl. Acad. Sci. U.S.A. **100**(3), 958–963.

Cooper, N. P. (**1999**). ''Vibration of beads placed on the basilar membrane in the basal turn of the cochlea,'' ARLO **106**(6), L59–L64.

Dallos, P. (**1992**). ''The active cochlea,'' J. Neurosci. **12**(12), 4575–4585.

Dimitriadis, E. K., and Chadwick, R. (**1999**). ''Solution of the inverse problem for a linear cochlear model: A tonotopic cochlear amplifier,'' J. Acoust. Soc. Am. **106**, 1880–1892.

Fettiplace, R., Ricci, A. J., and Hackney, C. M. (**2001**). ''Clues to the cochlear amplifier from the turtle ear,'' Trends Neurosci. **24**(3), 169–175.

Frank, G., Hemmert, W., and Gummer, A. W. (**1996**). ''Limiting dynamics of high-frequency electromechanical transduction of outer hair cells.'' Proc. Natl. Acad. Sci. **96**(8), 4410–4425.

Gummer, A. W., Hemmert, W. and Zenner, H. P. (**1996**). ''Resonant tectorial membrane motion in the inner ear: its crucial role in frequency tuning.'' Proc. Natl. Acad. Sci **93**(16), 8727–8732.

Hubbard, A. E., and Mountain, D. C. (**1983**). ''Alternating current delivered into the scala media alters sound pressure at the eardrum,'' Science **222**(4623), 510–512.

Kakehata, S., and Santos-Sacchi, J. (**1996**). ''Effects of salicylate and lanthanides on outer hair cell motility and associated gating charge,'' J. Neurosci. **16**(16), 4881–4889.

Khanna, S. M., Ulfendahl, M., and Steele, C. R. (**1998**). ''Vibration of reflective beads placed on the basilar membrane,'' Hear. Res. **116**, 71–85.

Kirk, D. L., and Yates, G. K. (**1996**). ''Frequency tuning and acoustic enhancement of electrically evoked otoacoustic emissions in the guinea pig cochlea,'' J. Acoust. Soc. Am. **100**(6), 3714–3725.

Liberman, M. C., Gao, J. G., He, D. Z. Z., Wu, X. D., Jia, S. P., and Zuo, J. (**2002**). ''Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier,'' Nature (London) **419**(6904), 300–304.

Ludwig, J., Oliver, D., Frank, G., Klocker, N., Gummer, A. W., and Falker, B. (**2001**). ''Reciprocal electromechanical properties of rat prestin: The motor molecule from rat outer hair cells,'' Proc. Natl. Acad. Sci. U.S.A. **98**(7), 4178–4183.

Nilsen, K. E., and Russell, I. J. (**2000**). ''The spatial and temporal representation of a tone on the guinea pig basilar membrane,'' Proc. Natl. Acad. Sci. U.S.A. **97**(22), 11751–11758.

Nuttall, A. L. (**1995**). ''Electromotile hearing: Evidence from basilar-membrane motion and otoacoustic emissions,'' Hear. Res. **92**, 170–177.

Nuttall, A. L., and Dolan, D. F. (**1996**). ''Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig,'' J. Acoust. Soc. Am. **99**(3), 1556–1565.

Nuttall, A. L., Zheng, J., Ren, T., and de Boer, E. (**2001**). ''Electrically evoked otoacoustic emissions from apical and basal perilymphatic electrode positions in the guinea pig cochlea,'' Hear. Res. **152**(1–2), 77–89.

Parthasarathi, A. A., Grosh, K., Nuttall, A. L., and Zheng, J. (**2003**). ''Influence of direct current stimulation on the *in vivo* basilar-membrane velocity response,'' J. Acoust. Soc. Am. **114**, 442–452.

Petrov, A. G., and Sachs, F. (**2002**). ''Flexoelectricity and elasticity of asymmetric biomembranes,'' Phys. Rev. E **65**, 021905.
Stewart, C. E., and Hudspeth, A. J. (**2000**). ''Effects of salicylates and aminoglycosides on spontaneous otoacoustic emissions in the Tokay gecko,'' Proc. Natl. Acad. Sci. U.S.A. **97**(1), 454–459.
Tunstall, M. J., Gale, J. E., and Ashmore, J. (**1995**). ''Action of salicylate on membrane capacitance of outer hair-cells from the guinea-pig cochlea,'' J. Physiol. (London) **485**(3), 739–752.
Weitzel, E. K., Tasker, R., and Brownell, W. E. (**2003**). ''Outer hair cell piezoelectricity: Frequency response enhancement and resonance behavior,'' J. Acoust. Soc. Am. **114**(3), 1462–1466.
Xue, S. (**1996**). ''Scala media voltage responses to sinusoidal current stimulation,'' *International Symposium on Diversity in Auditory Mechanics, University of California-Berkeley* (World Scientific, Singapore).
Xue, S., Mountain, D. C., and Hubbard, A. (**1995**). ''Electrically evoked basilar-membrane motion,'' J. Acoust. Soc. Am. **97**(5, Pt 1), 3030–3041.

Integration of outer hair cell activity in a one-dimensional cochlear model

Azaria Cohen and Miriam Furst[a)]
School of Electrical Engineering, Faculty of Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel

(Received 2 February 2004; revised 18 February 2004; accepted 19 February 2004)

Recently, significant progress has been made in understanding the contribution of the mammalian cochlear outer hair cells (OHCs) to normal auditory signal processing. In the present paper an outer hair cell model is incorporated in a complete, time-domain, one-dimensional cochlear model. The two models control each other through cochlear partition movement and pressure. An OHC gain (γ) is defined to indicate the outer hair cell contribution at each location along the cochlear partition. Its value ranges from 0 to 1: $\gamma=0$ represents a cochlea with no active OHCs, $\gamma=1$ represents a nonrealistic cochlea that becomes unstable at resonance frequencies, and $\gamma=0.5$ represents an ideal cochlea. The model simulations reveal typical normal and abnormal excitation patterns according to the value of γ. The model output is used to estimate normal and hearing-impairment audiograms. High frequency loss is predicted by the model, when the OHC gain is relatively small at the basal part of the cochlear partition. The model predicts phonal trauma audiograms, when the OHC gain is random along the cochlear partition. A maximum threshold shift of about 60 dB is obtained at 4 kHz. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1699391]

PACS numbers: 43.64.Kc, 43.64.Wn, 43.64.Ld [WPS] Pages: 2185–2192

I. INTRODUCTION

In recent years, significant progress has been made in understanding the contribution of the mammalian cochlear outer hair cells to normal auditory signal processing. The outer hair cells (OHCs) act as local amplifiers that are metabolically activated, and their motion effectively changes the basilar membrane mechanical response (e.g., Dallos, 1997). Hearing-impairment due to acoustic trauma or aging is mainly caused by outer hair cells loss (Moore, 1998).

In models that describe the outer hair cell activity, it is assumed that the OHC cilia displacement generates a force that acts on the basilar membrane (e.g., Allen and Neely, 1992; Mountain and Hubbard, 1994; Dallos and Evans, 1995; Rattay *et al.*, 1998; He and Dallos, 2000; Spector, 2000). The contribution of the OHCs to psychoacoustical performance has been previously demonstrated by models that functionally include OHC activity. These models predict normal behavior and degradation in the performances due to OHC loss (e.g., Kates, 1991; Goldstein, 1993; Carney, 1994; Heinz *et al.*, 2001).

Classical one-dimensional cochlear models have been modified to include the OHC activity by nonlinear damping and/or nonlinear compliance with additional delays (Talmadge *et al.*, 1998; Zweig, 1991). These models predicted various properties associated with OHCs, including cochlear otoacoustic emissions (Furst *et al.*, 1992; Talmadge *et al.*, 1998) and two-tone-suppression (e.g., Geisler, 1991, 1993; Neely, 1993).

In the present paper, we develop an integrated basilar membrane–outer hair cell model (BM-OHC). An outer hair cell model is embedded in a complete, time-domain, one-dimensional cochlear model. The two models control each other through the cochlear partition's movement and pressure. From the BM-OHC model, we derive predictions regarding normal hearing and various types of hearing impairments.

II. MODEL FORMULATION

A. Cochlear fluid dynamics

In the simple, one-dimensional model (e.g., Zwislocki, 1950; Zweig *et al.*, 1976; Viergever, 1980; Furst and Goldstein, 1982), the cochlea is considered as an uncoiled structure with two fluid-filled compartments whose walls, separated by an elastic partition, are rigid. In the present model we include the OHC activity in the elastic partition. By applying fundamental physical principles, such as conservation of mass and the dynamics of deformable bodies, we obtain the basic equations for this activity.

We define P_T and P_V as the pressure in scala tympani and scala vestibuli, respectively. The fluid velocity in the scala tympani is defined as U_T, and in the scala vestibuli as U_V. The pressure difference across the cochlear partition is

$$P=P_T-P_V. \tag{1}$$

Since both scalae, tympani and vestibuli, contain perilymph, which is presumably very nearly incompressible, and inviscid fluid, the equation of motion for each scala can be written as

$$\rho\frac{\partial U_T}{\partial t}=-\frac{\partial P_T}{\partial x},\quad \rho\frac{\partial U_V}{\partial t}=-\frac{\partial P_V}{\partial x}, \tag{2}$$

where ρ is the perilymph density.

The cochlear partition, whose mechanical properties are describable in terms of point-wise mass density, stiffness and damping, is regarded as a flexible boundary between scala

[a)]Author to whom correspondence should be addressed. Electronic mail: mira@eng.tau.ac.il

tympani and scala vestibuli. Thus at every point along the cochlear duct, the pressure difference (P) across the partition drives the partition's velocity. From the conservation of mass principle we can derive the relation between the fluid velocity and the basilar membrane displacement, ξ_{BM},

$$A\frac{\partial U_V}{\partial x}=\beta\frac{\partial \xi_{BM}}{\partial t},\quad A\frac{\partial U_T}{\partial x}=-\beta\frac{\partial \xi_{BM}}{\partial t}, \tag{3}$$

where A represents the cross-sectional area of scala tympani and scala vestibuli, and β is the basilar membrane width. Combining Eqs. (1)–(3) yields the differential equation for P,

$$\frac{\partial^2 P}{\partial x^2}=\frac{2\rho\beta}{A}\frac{\partial^2 \xi_{BM}}{\partial t^2}. \tag{4}$$

If we assume that the oval and round window areas are equal to A, then the boundary conditions are

$$U_V(0,t)=-U_T(0,t)=f(t),\quad P(L,t)=0, \tag{5}$$

where $f(t)$ is the stapes velocity, and L is the cochlear length.

In previous one-dimensional cochlear models (e.g., Zwislocki 1950; Viergever, 1980; Furst and Goldstein, 1982), the pressure difference is only obtained by the basilar membrane mechanical properties. In the present model, we include the pressure produced by the OHCs, therefore

$$P_{BM}=P+P_{OHC}, \tag{6}$$

where

$$P_{BM}=m(x)\cdot\frac{\partial^2 \xi_{BM}}{\partial t^2}+r(x)\cdot\frac{\partial \xi_{BM}}{\partial t}+s(x)\cdot\xi_{BM}, \tag{7}$$

and m, r, and s represent basilar membrane mass, resistance, and stiffness per unit area, respectively.

The next section presents the model for the outer-hair cell and the corresponding equations for P_{OHC}.

B. The outer-hair cell model

The outer-hair cells (OHC) are epithelial cells that are located in two different electrochemical environments. The apical part of the cell faces the endolymph of the scala media, while the basolateral part faces the supporting cells and the perilymph of the scala tympani. The OHC's stereocilia are bent due to basilar membrane motion; because their tip links are stretched and physically open ion channels, potassium and calcium ions flow from the endolymph into the cells. As a result, an electrical voltage drop is introduced on the basolateral of the cell's membrane (Dallos, 1997). These electrical changes change the shape of the OHCs (Jerry *et al.*, 1995; Brownell *et al.*, 1985). The somatic elongation-contraction process operates on a cycle-by-cycle basis at the frequency stimulus and provides a mechanical feedback to the basilar membrane motion. This process underlies the high sensitivity and frequency selectivity of the mammalian cochlea. The outer hair cell membrane acts as a low-pass filter with a cutoff frequency of less than 1 kHz. Dallos and Evans (1995) proposed that the extracellular potential gradient across the hair cell provides electro-motility at high frequencies. The following model is a modification of this proposal.

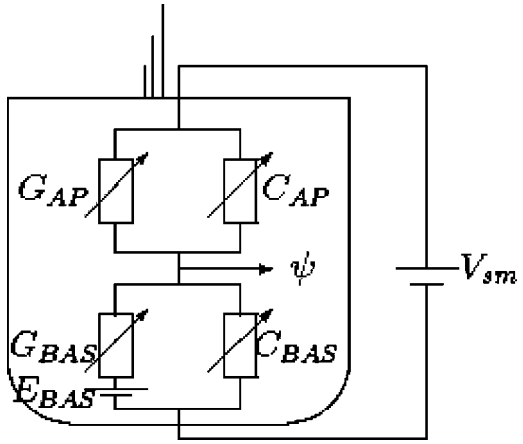


FIG. 1. An equivalent electrical circuit of the outer-hair cell model.

Figure 1 presents an equivalent electrical circuit model for the OHC. The cell membrane is divided into two parts, the apical and basolateral, each represented by its conductance (G_{AP} or G_{BAS}), and capacitance (C_{AP} or C_{BAS}). We assume that the stereocilia movement affects both the capacitance and conductance of the OHC's apical part. The electrical potential across the cell membrane is presumably equal to the potential of the scala media, V_{SM}, and the potential difference across the basolateral part (the receptor potential) is presented by ψ in Fig. 1. A battery (E_{BAS}), in series with the basolateral conductance, sets the resting OHC potential at roughly -70 mV (Mountain and Hubbard, 1995). The stereocilia motion causes a current to flow through the apical part of the OHC which yields apical current I_{AP},

$$I_{AP}=(V_{SM}-\psi)\cdot G_{AP}+\frac{d}{dt}[C_{AP}\cdot(V_{SM}-\psi)]. \tag{8}$$

The basolateral current I_{BAS}, on the other hand, is given by

$$I_{BAS}=G_{BAS}\cdot(\psi-E_{BAS})+\frac{d}{dt}[C_{BAS}\cdot\psi]. \tag{9}$$

According to Kirchhoff's current law, the current that flows through the apical segment is equal to the current that flows through the basolateral segment. Thus $I_{AP}=I_{BAS}$ yields the differential equation for ψ,

$$(C_{BAS}+C_{AP})\cdot\frac{d\psi}{dt}+\left(G_{AP}+G_{BAS}+\frac{dC_{BAS}}{dt}+\frac{dC_{AP}}{dt}\right)\cdot\psi$$
$$=V_{SM}\cdot\left(G_{AP}+\frac{dC_{AP}}{dt}\right)+C_{AP}\frac{dV_{SM}}{dt}.+G_{BAS}\cdot E_{BAS}. \tag{10}$$

In order to simplify Eq. (10), we apply the following assumptions:

(i) The capacitance and the conductance of the OHC membrane are proportional to the membrane surface area. The basolateral membrane area is larger than the apical part. According to Housley and Ashmore (1992) the ratio between

the basolateral and apical capacitance is more than 16, and the ratio of the conductance is more than 7. We thus assume $C_{AP} \ll C_{BAS}$ and $G_{AP} \ll G_{BAS}$.

(ii) The changes in the basolateral capacitance are relatively small (Lukashkin and Russell, 1997), thus

$$\frac{dC_{BAS}}{dt} \approx 0. \tag{11}$$

(iii) We shall further assume that $|dC_{AP}/dt|$ is at most in the order-of-magnitude of $|G_{AP}|$. Thus from assumption (i) we can conclude that

$$\left|\frac{dC_{AP}}{dt}\right| \ll |G_{BAS}|. \tag{12}$$

(iv) The OHC membrane acts as a low pass filter with a fixed cutoff frequency of less than 1 kHz (Dallos and Evans, 1995). The OHC cutoff frequency (ω_{OHC}) can be derived from

$$\omega_{OHC} = \frac{G_{AP}+G_{BAS}}{C_{AP}+C_{BAS}} \approx \frac{G_{BAS}}{C_{BAS}} = \text{const.} \tag{13}$$

(v) The electrical potential across the scala media is presumably constant (Mountain and Hubbard, 1995), thus

$$\frac{dV_{SM}}{dt} \approx 0. \tag{14}$$

Substituting Eqs. (11)–(14) in Eq. (10) yields a simplified differential equation for ψ:

$$\frac{d\psi}{dt} + \omega_{OHC} \cdot [\psi - E_{BAS}] = \lambda \cdot \left[G_{AP} + \frac{dC_{AP}}{dt} \right], \tag{15}$$

where $\lambda = V_{SM}/[C_{BAS}+C_{AP}] \approx V_{SM}/C_{BAS} = \text{const.}$

The apical membrane's capacitance and conductance change due to ion channels actively opening in the apical part of the outer hair cell. The OHC stereocilia are shallowly but firmly embedded in the under-surface of the tectorial membrane. Since this membrane is attached on one side to the basilar membrane, a shear motion arises between the tectorial membrane and the organ of Corti, when the basilar membrane moves up and down (Pickles, 1988). The arch of Corti maintains the rigidity of the organ during the movement.

As part of our one-dimensional assumption, we further assume that the tectorial membrane along with Dieters cells is a rigid structure. Thus it is reasonable to assume that the OHC bundle displacement is a function of the basilar membrane vertical displacement (Nobili *et al.*, 1998). We, therefore, assume that both G_{AP} and C_{AP} are functions of ξ_{BM} (the basilar membrane vertical displacement). Their functional dependences are best described by second order Boltzmann functions (e.g., Sachs and Lecar, 1991; Lukashkin and Russel, 1998).

The voltage variation across the basolateral part of the OHC causes a length change (Δl_{OHC}) in the OHC as Brownell *et al.* (1985) first showed. Thus the force F_{OHC} that an OHC exhibits due to voltage change is derived by

$$F_{OHC} = K_{OHC} \cdot [\xi_{BM} + \Delta l_{OHC}], \tag{16}$$

where $\Delta l_{OHC} = \Delta l_{OHC}(\psi(\xi_{BM}))$. The pressure that the OHCs contribute to the basilar membrane pressure is derived from

$$P_{OHC} = \gamma(x) \cdot F_{OHC}, \tag{17}$$

where $\gamma(x)$ is the relative density of healthy OHCs per unit area along the cochlear duct. In the rest of the paper we will refer to $\gamma(x)$ as the OHC gain, whose value ranges from 0 to 1.

C. Linear approximation

The contribution of the OHCs to the basilar membrane motion is particularly effective in low level stimuli. Therefore, the second order Boltzmann functions can be approximated by linear functions for the capacitance, conductance, length change, and OHC stiffness.

In the following section we shall assume linear dependence, i.e.,

$$\begin{aligned} &G_{AP} = \alpha_G^0 + \alpha_G \cdot \xi_{BM}, \quad C_{AP} = \alpha_C^0 + \alpha_C \cdot \xi_{BM}, \\ &\Delta l_{OHC} = \alpha_l^0 - \alpha_l \cdot \psi, \end{aligned} \tag{18}$$

where α_G, α_C, and α_l are all positive numbers. Substituting the linear assumptions (18) in Eqs. (15)–(17) yields the differential equation for P_{OHC}:

$$\frac{\partial P_{OHC}}{\partial t} + \omega_{OHC} \cdot P_{OHC} = \gamma \cdot \left[\alpha_1 \cdot \xi_{BM} + \alpha_2 \cdot \frac{\partial \xi_{BM}}{\partial t} + \alpha_0 \right], \tag{19}$$

where

$$\begin{aligned} &\alpha_0 = K_{OHC} \cdot [\omega_{OHC} \cdot (\alpha_l^0 - \alpha_l \cdot E_{BAS}) - \lambda \cdot \alpha_l \cdot \alpha_G^0], \\ &\alpha_1 = K_{OHC} \cdot (\omega_{OHC} - \lambda \cdot \alpha_l \cdot \alpha_G), \\ &\alpha_2 = K_{OHC} \cdot (1 - \lambda \cdot \alpha_l \cdot \alpha_C). \end{aligned} \tag{20}$$

When $\gamma = 0$, the OHCs are not contributing to the basilar membrane motion, and thus $P_{OHC} = 0$. On the other hand, $\gamma = 1$ means that the basilar membrane displacement can reach infinity at the location whose characteristic frequency is equal to the stimulus frequency.

In order to solve Eq. (19), all the parameters in Eq. (20) should be determined. Without losing the generality, we can assume that $\alpha_0 = 0$ [this term contributes to the nonhomogeneous part of Eq. (19) which can be accounted by the initial conditions]. The values of α_1 and α_2 were obtained from the frequency domain solution.

For a steady-state sinusoidal input, the frequency domain representation of Eq. (19) is given by

$$P_{OHC}(\omega, x) = \gamma(x) \cdot \frac{\alpha_1(x) + j\omega \cdot \alpha_2(x)}{\omega_{OHC} + j\omega} \xi_{BM}(\omega, x). \tag{21}$$

Substituting Eq. (21) in the frequency domain representation of Eqs. (6) and (7) yields

$$P(x, \omega) = j \cdot \omega \cdot Z(x, w) \cdot \xi_{BM}(x, \omega), \tag{22}$$

where

$$\mathrm{Re}\{Z(x,\omega)\}=r(x)-\gamma(x)\frac{\alpha_2(x)\cdot\omega_{OHC}-\alpha_1(x)}{\omega^2+\omega_{OHC}^2},$$
$$\mathrm{Im}\{Z(x,\omega)\}=m(x)\cdot\omega-\frac{s(x)}{\omega}+\frac{\gamma(x)}{\omega}\cdot\frac{\alpha_1(x)\cdot\omega_{OHC}+\alpha_2(x)\cdot\omega^2}{\omega^2+\omega_{OHC}^2}. \quad (23)$$

We further assume that for every location along the cochlear partition, when the stimulus frequency (ω) is equal to $\omega_{CF}(x)=\sqrt{s(x)/m(x)}$, a resonance is produced. The resonance frequency in a cochlea with no active OHCs [$\gamma(x)\equiv 0$] was obtained when $\mathrm{Im}\{Z(x,\omega_{CF})\}=0$ but $\mathrm{Re}\{Z(x,\omega_{CF})\}\neq 0$. When $\gamma(x)\equiv 1$, we require that at the resonance frequency both $\mathrm{Re}\{Z(x,\omega_{CF})\}=0$ and $\mathrm{Im}\{Z(x,\omega_{CF})\}=0$. Substituting those requirements in Eq. (23) reveals the values for $\alpha_1(x)$ and $\alpha_2(x)$,

$$\alpha_1(x)=-\frac{r(x)\cdot s(x)}{m(x)},\quad \alpha_2(x)=r(x)\cdot\omega_{OHC}. \quad (24)$$

Comparison between Eqs. (24) and (20) yields some constraints on the parameters α_G, α_C, and α_l; specifically, $\lambda\cdot\alpha_l\cdot\alpha_G>\omega_{OHC}$ and $\lambda\cdot\alpha_l\cdot\alpha_C<1$, which implies that $\alpha_C<\alpha_G/\omega_{OHC}$. This result might explain why changes in the OHC's apical conductance were more frequently detected than its capacitance changes.

III. SIMULATION METHODS

A. Frequency domain solution

For low-level, steady-state stimuli, we consider a linear model that can be solved in the frequency domain. The model Eq. (4) for the pressure difference, $P(x,\omega)$, is written in the frequency domain as

$$\frac{d^2P(x,\omega)}{dx^2}+K^2(x,\omega)\cdot P(x,\omega)=0, \quad (25)$$

where $K^2(x,\omega)=2j\omega\rho\beta/AZ(x,\omega)$; $Z(x,\omega)$ is obtained by substituting Eq. (24) in Eq. (23). The frequency domain boundary conditions are obtained by transforming Eq. (5) to the frequency domain and substituting the result in the frequency domain representation of Eq. (2), which yields

$$\left.\frac{dP(x,\omega)}{dx}\right|_{x=0}=2j\omega\rho F(\omega),\quad P(L,\omega)=0, \quad (26)$$

where $F(\omega)$ is Fourier transform of $f(t)$.

The most common method used for solving the cochlear boundary value problem [Eqs. (25) and (26)] is the WKB method (e.g., Viergever, 1980). In this approximation, one assumes that $K(x)$ is varying slowly with x. However, near CF, this assumption does not hold, especially when we assume a significant existence of OHC population [$\gamma(x)\rightarrow 1$]. We chose to solve the boundary value problem using a finite difference method, one of the simplest ways to solve these sort of problems (e.g., Burden and Faires, 1997). Although more sophisticated methods such as finite elements might provide better approximations to the boundary value problem, the finite difference method is significantly more efficient for a dense net (Diependaal, 1987). A detailed description of the frequency domain solution is included in Appendix A.

TABLE I. List of model parameters.

Parameter	Value/definition	Description
L	3.5 cm	Cochlear length
ρ	1 g/cm^3	Density of perilymph
β	0.003 cm	Width of the basilar membrane
A	0.5 cm^2	Cross-sectional area of the cochlea scalae
m	$1.286\times10^{-6}\cdot e^{1.5\cdot X}$ g/cm^2	Basilar membrane mass per unit area
s	$1.282\times10^{4}\cdot e^{-1.5\cdot X}$ g/cm$^2\cdot$s^2	Basilar membrane stiffness per unit area
r	$0.25\times e^{-0.06\cdot X}$ g/cm$^2\cdot$s	Basilar membrane resistance per unit area
ω_{OHC}	1 kHz	OHC cutoff frequency
x	cm	Distance from stapes
$f(t)$	cm/s	Stapes velocity

B. Time domain solution

The time domain solution is performed in two stages (Furst and Goldstein, 1982). In the first stage, the boundary value differential equation is solved when the time is held as a parameter. In the second stage, the set of the initial condition time dependent equations is solved.

In order to solve the boundary value differential equation [Eq. (4)], we rewrite the equation by substituting Eqs. (6) and (7) in Eq. (4), which yields

$$\frac{\partial^2P}{\partial x^2}-\Omega(x)\cdot P=g(x,t)\cdot\Omega(x), \quad (27)$$

where $\Omega(x)=2\rho\beta/mA$, and $g(x,t)=-[r\,\partial\xi_{BM}/\partial t+s\xi_{BM}]+P_{OHC}$.

The initial value differential equations to be solved are Eqs. (6), (7) and (19) with the initial conditions

$$\xi_{BM}(x,t=0)=0,\quad \left.\frac{\partial\xi_{BM}(x,t)}{\partial t}\right|_{t=0}=0,$$
$$P_{OHC}(x,t=0)=0. \quad (28)$$

The actual simulation method is described in Appendix B. The parameters that were used in the simulation are listed in Table I.

IV. SIMULATIONS RESULTS

Figure 2 represents relative basilar membrane velocities (BM_V) along the cochlear length as obtained by various sinusoidal inputs. Both time and frequency solutions yielded similar results for those steady-state conditions. Each set of BM_V was derived with a different value of γ. In these calculations, a fixed γ was considered along the whole cochlear

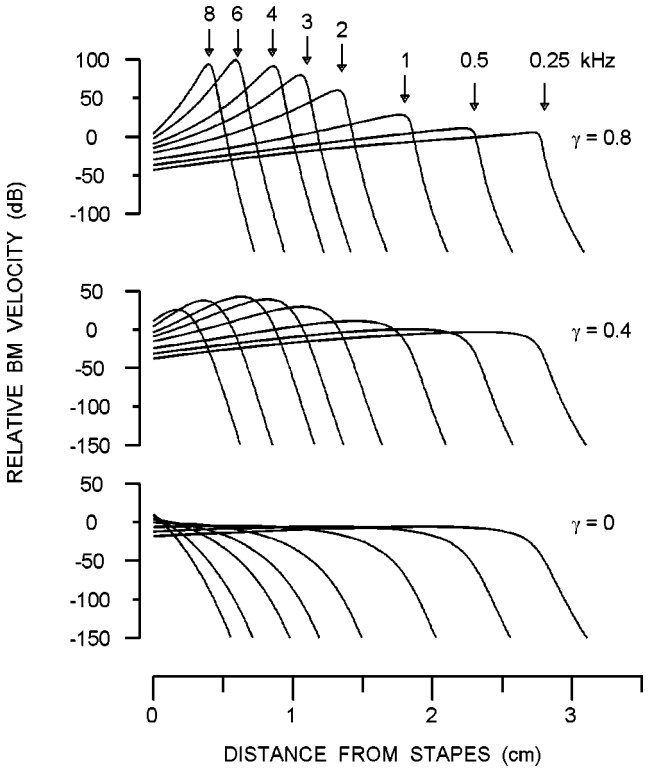


FIG. 2. Simulation of relative basilar membrane velocity for different input frequencies and various values of γ, the OHC gain.

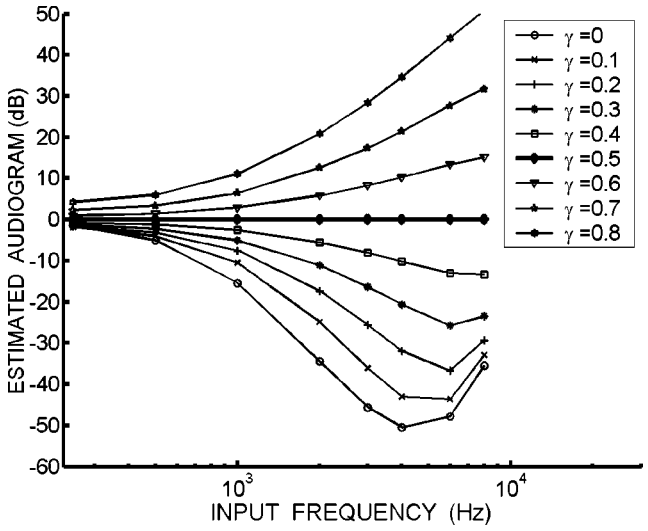


FIG. 3. Simulated audiograms derived by various values of γ.

partition. It is obvious from Fig. 2 that with the increase of γ, the location of the resonance for each input frequency moves towards the helicotrema, and the peak at the resonance becomes more significant, especially for frequencies above 1 kHz.

We chose $\gamma(x)=0.5$ to represent the ideal cochlea. Let us define loudness (L_d) as the energy acquired by the whole cochlea due to the basilar membrane velocity (Furst *et al.*, 1992), i.e.,

$$L_d=\frac{1}{T}\int_0^L\int_0^T\left[\frac{\partial\xi_{BM}}{\partial t}\right]^2 dt\cdot dx, \tag{29}$$

where T is the stimulus duration. Human cochlear activity is best described by its audiogram, which is the threshold obtained in different stimulus frequencies relative to an ideal cochlea. Thus we can use the model simulation in order to estimate the audiograms that will be obtained with different values of γ. Let us define $\Theta(\omega,\gamma)$ as the threshold obtained by the model for an input frequency ω and a cochlea with an OHC gain γ,

$$\Theta(\omega,\gamma)=L_d(\omega,0.5)-L_d(\omega,\gamma). \tag{30}$$

It is obvious then from Eq. (30) that $\Theta(\omega,0.5)=0$, for every ω. Figure 3 represents the estimated audiograms for different values of γ. For cochleae with $\gamma>0.5$, the estimated audiograms reveal thresholds that are better than that for the ideal cochlea. The difference in the estimated thresholds for different values of γ is less than 20 dB for input frequencies below 1 kHz. However, there is a significant difference in the estimated threshold for higher frequencies. For $\gamma<0.2$, each of the estimated audiograms has a maximum threshold at a frequency between 4 and 6 kHz. Those types of audiograms resemble typical phonal trauma audiograms. Higher values of γ ($\gamma>0.5$), however, reveal audiograms with decreased threshold as a function of frequency. These types of audiograms are not realistic, since they are not found among human audiograms. We, therefore, define a normal healthy cochlea as one whose OHC gain (γ) is equal to 0.5, and a cochlea with OHC loss is presented by $\gamma<0.5$.

Most studies that measured OHC loss on humans or animals (e.g., Liberman and Dodds, 1987; Saunders *et al.*, 1985) demonstrated that OHC loss typically starts at the basal part of the cochlear partition and gradually expands toward the helicotrema. In order to test the model prediction for such a case, we assumed a gradual change in $\gamma(x)$ and calculated the corresponding audiograms. Figure 4(a) presents a set of OHC gains obtained by $\gamma(x)=0.5\cdot(1-e^{-\alpha x})$, for different values of α. The corresponding audiograms are presented in Fig. 4(b). The estimated audiograms resemble hearing-impaired human audiograms that are typically found among the elderly. While there is almost no hearing loss at low frequencies below 1 kHz, there is a gradual decrease in hearing for higher frequencies.

The OHC gain (γ) can also be interpreted as the density of the OHC population along the cochlear partition. It is reasonable to assume that this density is not fixed, but randomly varies along the cochlea. Let us assume that $\gamma(x)$ is a Gaussian random variable with a mean of 0.5 and a standard deviation σ. Figure 5 shows the excitation patterns obtained for different values of σ while the corresponding audiograms appear in Fig. 6. Each BM_V in Fig. 5 was obtained from 50 runs of the model; for each run, a different $\gamma(x)$ was generated. It is clear from Fig. 5 that low values of σ yielded very similar responses obtained by a fix γ as shown in Fig. 2. However, for larger values of σ, the different BM_V's lose their significant peaks and resemble the responses in Fig. 2 that were obtained for $\gamma=0$. Figure 6 presents the correspondent mean and standard deviation thresholds. An obvious maximum threshold was obtained at 4 kHz when $\sigma>0.01$. A maximum threshold of 60 dB was obtained for $\sigma=0.1$. The

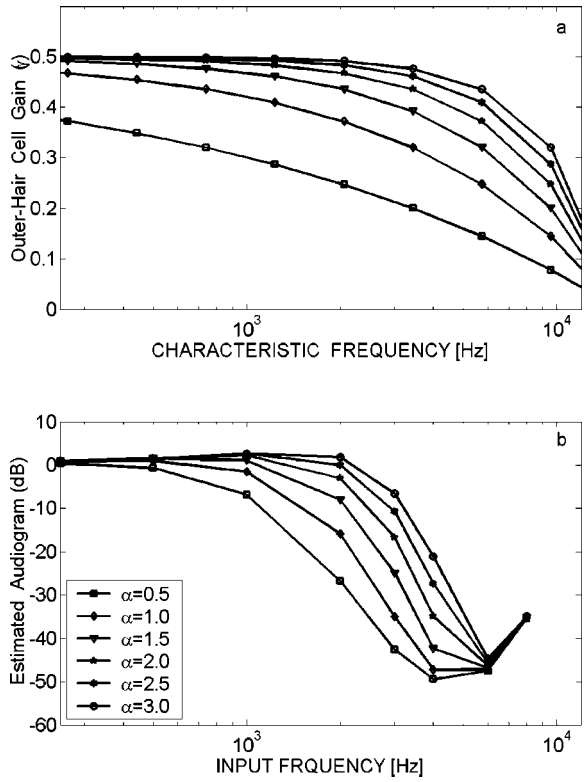


FIG. 4. Simulated audiograms (b) for different types of OHC hearing loss as characterized by the OHC gain (a).

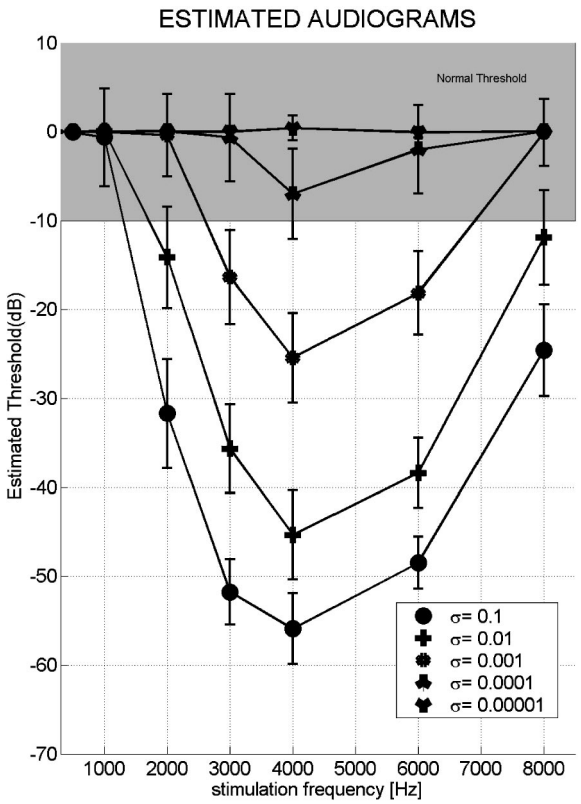


FIG. 6. Simulated audiograms obtained for normally distributed γ with a mean of 0.5 and different values of variance σ^2.

estimated audiograms resemble typical, noise-induced hearing loss audiograms.

V. DISCUSSION

The model presented in this paper is a modification of the classical one-dimensional cochlear model, where the OHC activity is integrated in the basilar membrane motion. The OHCs generate forces due to their electro-motility. These forces act simultaneously along with the trans-partition pressure that the cochlear fluids produce. In the vicinity of the characteristic frequency (CF), these forces are in phase for a stimulus at the frequency of the CF and yield a significant enhancement in the basilar membrane motion. The response to stimuli that are equal to CF is reduced, because these forces are not in phase.

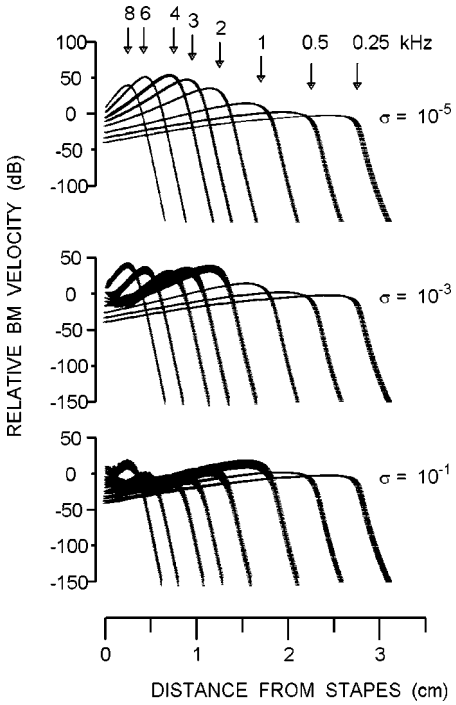


FIG. 5. Simulation of basilar membrane velocity for different input frequencies and normally distributed γ with a mean of 0.5 and different values of variance σ^2. Each BM_V was obtained from 50 runs; mean and standard deviation are shown.

The OHCs are known to be sensitive to metabolism, efferent stimulation, etc. (Ulfendahl, 1997). In the present model their mode of activity is characterized by one parameter γ, which can vary along the cochlear partition. We refer to a cochlea whose $\gamma(x)=0$ as a passive cochlea, i.e., it has no active OHCs. On the other hand, $\gamma(x)=1$ represents a nonrealistic cochlea whose basilar membrane motion is unstable. An ideal cochlea was defined as one with $\gamma(x)=0.5$.

The main effect of OHC activity is to enhance cochlear displacement and velocity in the region where the stimulus frequency is close to the local characteristic frequency. Moreover, this enhancement is mostly significant for frequencies above 1 kHz (the cutoff frequency of the OHC membrane). Any random change in the OHC gain, even a very fine one, causes a decrease in the enhanced basilar membrane motion. On the basis of the basilar membrane velocity, we have simulated audiograms for normal ears and ears with OHC loss. The simulation results revealed the following observations:

(i) Typical audiograms with high frequency loss were obtained when the OHC gain (γ) decreased at the basal part of the cochlear partition.
(ii) Phonal trauma typical audiograms were obtained when the OHC gain was random along the cochlear partition.
(iii) The maximum threshold shift that was obtained due to OHC loss was about 60 dB at 4 kHz.

Noise-induced hearing loss is typically presented by an audiogram whose maximum threshold is obtained at 4 kHz, similar to the simulated audiograms with random γ. The model therefore indicates that random loss of OHCs along the entire cochlear partition characterizes phonal trauma. This description of noise-induced cochleae is incompatible with animal studies on noise trauma. These studies (Liberman and Mulroy, 1982; Maison and Liberman, 2000) show that among animals exposed to narrow-band noise signals, the OHC loss was especially found at cochlear locations whose characteristic frequencies were equal to or an octave above the noise exposure. Such simulation of OHCs loss will also produce similar audiograms. However, the assumption of random loss of OHCs along the cochlea seems more realistic, since in humans the loss of sensitivity in the 4-kHz range was found independently of the type of noise exposure (Saunders *et al.*, 1985; Moore, 1998).

Hearing-impairment due to advanced age is typically characterized by high frequency hearing loss. The typical audiogram for this sort of hearing loss is predicted by the model when OHC loss is considered to be primarily occurring at the basal part of the cochlea. Indeed, studies on cadaver cochlea indicate that that OHC loss in the basal part is more frequent than in the apical part of the cochlea (e.g., Saunders *et al.*, 1985).

The model predicts a maximum of 60 dB hearing loss, when in practice hearing loss can be as high as 120 dB. Since the model deals only with OHC loss, it is very reasonable to assume that loss of inner hair-cells that typically follow OHC loss are the cause of the profound hearing loss.

The time-domain solution presented in this paper allows us to test the model's predictions for known nonlinear phenomena such as two-tone suppression, combination tones, and cochlear otoacoustic emissions. Hearing-impaired people who suffer from OHC loss show a significant degradation in these nonlinear phenomena (Moore, 1998). In future studies, we intend to demonstrate the model's predictions for these nonlinear phenomena.

ACKNOWLEDGMENTS

We thank two unnamed reviewers whose comprehensive and detailed comments were very helpful.

APPENDIX A: FREQUENCY DOMAIN SOLUTION

Let us define a uniform grid of N points in the interval $[0,L]$, so that $x_l = lh$ and $P_l = P(x_l,\omega)$, where $h = L/N$, and $l = 0,1,...,N-1$. Applying linear combinations of the Taylor expansions for P_{l-1} and P_{l+1} yields

$$\left.\frac{d^2P}{dx^2}\right|_{x=x_l} = \frac{P_{l-1}-2P_l+P_{l+1}}{h^2}+o(h^2), \quad 1<l<N-2$$
$$\left.\frac{d^2P}{dx^2}\right|_{x=0} = \frac{2(P_1-P_0-hP_0')}{h^2}+o(h), \tag{A1}$$

where $P_0' = 2j\omega\rho\cdot F(\omega)$ is derived from the boundary values [Eq. (26)]. Note that for $1<l<N-2$, the derivatives of P are approximated by the order of h^2, while at the border point, $l=0$, P is approximated only by the order of h. Since we assume that P is a smooth function, $o(h)$ approximation at the edges might be sufficient. Thus Eq. (A1) can be expressed as a set of linear equations:

$$\Lambda\cdot\underline{P} = \underline{\Gamma}, \tag{A2}$$

where

$$\underline{P} = [P_0,P_1,\ldots,P_{N-1}]^T,$$

and

$$\underline{\Gamma} = 2hj\omega\rho\cdot F(\omega)\cdot[1,0,...,0]^T.$$

The matrix Λ is a tridiagonal square matrix of the size $N\times N$, whose values on the diagonal (Λ_{ll}) are obtained by

$$\Lambda_{ll} = \left\{\begin{array}{cc} -1+h^2\cdot K^2(x_0)/2, & l=0, \\ -2+h^2\cdot K^2(x_l), & 1\leqslant l\leqslant N-2, \\ 1, & l=N-1, \end{array}\right\}, \tag{A3}$$

and whose bordering values are $\Lambda_{l,l+1}=1$ for $l=0,...,N-2$, $\Lambda_{l,l-1}=1$ for $l=1,...,N-2$, and $\Lambda_{N-1,N-2}=0$. If the matrix Λ is regular, then a unique solution exists and can be easily obtained. Although we were not able to prove that for any choice of h, $\det(\Lambda)\neq 0$, in practice we have always been able to calculate Λ^{-1}.

APPENDIX B: TIME DOMAIN SOLUTION

The time domain solution is performed in two stages (Furst and Goldstein, 1982). In the first stage, the boundary value differential equation is solved when the time is held as a parameter. In the second stage, the set of the initial condition time dependent equations is solved.

The boundary-value problem [Eq. (27)] can be expressed as a set of linear equations

$$\Delta\cdot\underline{P} = \underline{Y}, \tag{B1}$$

where $\underline{P} = [P_0,P_1,\ldots,P_{N-1}]^T$ is the same as in the frequency domain solution, but it is obtained for every time step t. The boundary values Eq. (5) are expressed as $P_0' = 2\rho f'(t)$ and $P_{N-1}=0$, thus

$$\underline{Y} = \begin{bmatrix} 2h\rho f'(t)+\frac{1}{2}h^2g(x_0,t)\cdot\Omega(x_0) \\ h^2g(x_1,t)\cdot\Omega(x_1) \\ h^2g(x_{N-2},t)\cdot\Omega(x_{N-2}) \\ 0 \end{bmatrix}. \tag{B2}$$

The matrix Δ is a tridiagonal square matrix of the size $N\times N$, whose values on the diagonal (Δ_{ll}) are obtained by

$$\Delta_{ll}=\begin{cases} -[1+h^2\Omega(x_0)/2], & l=0, \\ -[2+h^2\Omega(x_l)], & 1\leqslant l\leqslant N-2, \\ 1, & l=N-1, \end{cases} \tag{B3}$$

and whose bordering values are $\Delta_{l,l+1}=1$ for $l=0,...,N-2$, $\Delta_{l,l-1}=1$ for $l=1,...,N-2$, and $\Delta_{N-1,N-2}=0$. Note that the matrix Δ is independent of t, thus there is no need to calculate Δ in every time step. For $t=0$, Eq. (B3) can be solved by using the initial conditions [Eq. (28)] and then calculating $g(x,t=0)$. For the next time step, $t=t+\Delta t$, ξ_{BM}, $\partial\xi_{BM}/\partial t$, and P_{OHC} are derived by using the modified Euler method (e.g., Iserles, 1997) for solving the set of differential equations (7) and (19).

Allen, J. B., and Neely, S. T. (**1992**). ''Micromechanical models of the cochlea,'' Phys. Today **July 1992**, 40–47.

Brownell, W. E., Bader, C. R., Bertrand, D., and de Ribaupierre, I. (**1985**). ''Evoked mechanical responses of isolated cochlear outer hair cells,'' Science **227**, 194–196.

Burden, R. L., and Faires, J. D. (**1997**). *Numerical Analysis* (Brooks-Cole, Belmont, MA).

Carney, L. H. (**1994**). ''Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness,'' Hear. Res. **76**, 31–44.

Dallos, P. (**1997**). ''Outer hair cells: The inside story,'' Ann Otol. Laryngol. **106**, 16–22.

Dallos, P., and Evans, B. N. (**1995**). ''High-frequency motility of outer hair-cells and the cochlear amplifier,'' Science **267**, 2006–2009.

Diependaal, R. J., Duifuis, H., Hoogstraten, H. W., and Viergeve, M. A. (**1987**). ''Numerical methods for solving one dimensional cochlear models in the time domain,'' J. Acoust. Soc. Am. **82**, 1655–1666.

Furst, M., and Goldstein, J. L. (**1982**). ''A cochlear transmission line model compatible with psychophysics,'' J. Acoust. Soc. Am. **72**, 717–726.

Furst, M., Reshef (Haran), I., and Attias, J. (**1992**). ''Manifestation of intense noise stimulation on spontaneous otoacoustic emission and threshold microstructure: Experiment and model,'' J. Acoust. Soc. Am. **91**, 1003–1014.

Geisler, C. D. (**1991**). ''A cochlear model using feedback from motile outer hair cells,'' Hear. Res. **54**, 105–117.

Geisler, C. D. (**1993**). ''A realizable cochlear model using feedback from motile outer hair cells,'' Hear. Res. **68**, 253–262.

Goldstein, J. L. (**1993**). ''Exploring new principles of cochlear operation: bandpass filtering by the organ of Corti and additive amplification in the basilar membrane,'' in *Biophysics of Hair Cell Sensory System*, edited by H. Duiffhuis, J. W. Horst, P. van Dijk, and S. M. van Netten (World Scientific, Singapore), pp. 315–322.

He, D. Z. Z., and Dallos, P. (**2000**). ''Properties of voltage dependant somatic stiffness of cochlear outer hair cell,'' JARO **1**, 64–81.

Heinz, M. G., Zhang, X., Bruce, I. C., and Carney, L. H. (**2001**). ''Auditory nerve model for predicting performance limits of normal and impaired listeners,'' JARO **2**, 91–96.

Housley, G. D., and Ashmore, J. F. (**1992**). ''Ionic currents of outer hair cells isolated from the guinea-pig cochlea,'' J. Physiol. **448**, 73–98.

Iserles, A. (**1997**). *A First Course in the Numerical Analysis of Differential Equations* (Cambridge U.P., Cambridge).

Jerry, R. A., Popel, A. S., and Brownell, W. E. (**1995**). ''Outer hair cell length changes in an external electric field. I. The role of intracellular electro-osmotically generated pressure gradients,'' J. Acoust. Soc. Am. **98**, 2000–2017.

Kates, J. M. (**1991**). ''A time domain digital cochlear model,'' IEEE Trans. Signal Process. **39**, 2573–2592.

Liberman, M. C., and Mulroy, M. J. (**1982**). ''Acute and chronic effects of acoustic trauma. Cochlear pathology and auditory nerve pathophysiology,'' in *New Perspectives on Noise-induced Hearing-loss*, edited by R. P. Hamernik, D. Henderson, and R. Salvi (Raven, New York), pp. 105–134.

Liberman, M. C., and Dodds, L. W. (**1987**). ''Acute ultrastructural changes in acoustic trauma: serial section reconstruction of sterocilia and cuticular plates,'' Hear. Res. **26**, 45–64.

Lukashkin, A. N., and Russell, I. J. (**1997**). ''The voltage dependence for the transducer may modify outer hair cell motility,'' Br. J. Audiol. **31**, 85–86.

Maison, S. F., and Liberman, M. C. (**2000**). ''Predicting vulnerability to acoustic injury with a noninvasive assay of olivocochlear reflex strength,'' J. Neurosci. **20**(12), 4701–4707.

Moore, B. C. J. (**1998**). *Cochlear Hearing Loss* (Academic, New York).

Mountain, D. C., and Hubbard, A. E. (**1994**). ''A piezoelectric model for outer hair cell function,'' J. Acoust. Soc. Am. **95**, 350–354.

Mountain, D. C., and Hubbard, A. E. (**1995**). ''Computational analysis of hair cell and auditory nerve processes,'' in *Auditory Computation*, edited by H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay, (Springer, New York), pp. 121–156.

Neely, S. T. (**1993**). ''A model of cochlear mechanics with OHC motility,'' J. Acoust. Soc. Am. **94**, 137–146.

Nobili, R., Mammano, F., and Ashmore, J. F. (**1998**). ''How well do we understand the cochlea?'' TINS **21**, 159–165.

Pickles, J. O. (**1988**). *An introduction to the physiology of hearing*, 2nd ed (Academic, New York).

Rattay, F., Gebeshuber, I. C., and Gitter, A. H. (**1998**). ''The mammalian auditory hair cell: A simple electric circuit model,'' J. Acoust. Soc. Am. **103**, 1558–1565.

Sachs, F., and Lecar, H. (**1991**). ''Stochastic models for mechanical transductions,'' Biophys. J. **59**, 1143–1145.

Saunders, J. C., Dear, S. P., and Schneider, M. E. (**1985**). ''The anatomical consequences of acoustic injury: A review and tutorial,'' J. Acoust. Soc. Am. **78**, 833–860.

Spector, A. A. (**2000**). ''On the mechanoelectrical coupling in the cochlear outer hair cell,'' J. Acoust. Soc. Am. **107**, 1435–1441.

Talmadge, C. L., Tubis, A., Long, G. L., and Piskorski, P. (**1998**). ''Modeling otoacoustic emission and hearing threshold fine structure,'' J. Acoust. Soc. Am. **104**, 1517–1543.

Ulfendahl, M. (**1997**). ''Mechanical responses of the mammalian cochlea,'' Prog. Neurobiol. **53**, 331–380.

Viergever, M. A. (**1980**). ''Mechanics of the inner ear a mathematical approach,'' Delft University of Technology, Netherlands.

Zweig, G. (**1991**). ''Finding the impedance of the organ of Corti,'' J. Acoust. Soc. Am. **89**, 1229–1254.

Zweig, G., Lipes, R., and Pierce, J. R. (**1976**). ''The cochlear compromise,'' J. Acoust. Soc. Am. **59**, 975–982.

Zwislocki, J. J. (**1950**). ''Theory of the acoustical action of the cochlea,'' J. Acoust. Soc. Am. **22**, 778–784.

Auditory steady-state responses reveal amplitude modulation gap detection thresholds

Bernhard Ross[a] and Christo Pantev
Institute for Biomagnetism and Biosignalanalysis, Münster University Hospital, University of Münster, Germany

(Received 19 October 2003; revised 28 January 2004; accepted 12 February 2004)

Auditory evoked magnetic fields were recorded from the left hemisphere of healthy subjects using a 37-channel magnetometer while stimulating the right ear with 40-Hz amplitude modulated (AM) tone-bursts with 500-Hz carrier frequency in order to study the time-courses of amplitude and phase of auditory steady-state responses (ASSRs). The stimulus duration of 300 ms and the duration of the silent periods (3–300 ms) between succeeding stimuli were chosen to address the question whether the time-course of the ASSR can reflect both temporal integration and temporal resolution in the central auditory processing. Long lasting perturbations of the ASSR were found after gaps in the AM sound, even for gaps of short duration. These were interpreted as evidences for an auditory reset mechanism. Concomitant psycho-acoustical tests corroborated that gap durations perturbing the ASSR were in the same range as the threshold for AM gap detection. Magnetic source localizations estimated the ASSR sources in the primary auditory cortex, suggesting that the processing of temporal structures in the sound is performed at or below the cortical level. [DOI: 10.1121/1.1694996]

Pages: 2193–2206

PACS numbers: 43.64.Ri, 43.64.Qh, 43.64.Bt [WPS]

I. INTRODUCTION

Humans can detect short discontinuities in sound of only a few milliseconds duration. However, the gap detection performance strongly depends upon the spectral composition of the stimulus and other experimental conditions like the duration of the leading and trailing sounds. In narrow-band noise the gap detection threshold (GT) increases with decreasing center frequency from about 5 ms at 8 kHz to 50 ms at 250 Hz (Buus and Florentine, 1985). In white noise GTs of 2–3 ms were reported by several studies (Buus and Florentine, 1985; Eddins *et al.*, 1992; Green and Forrest, 1989; Shailer and Moore, 1983). At least in narrow-band signals the GT decreases with increasing stimulus bandwidth (Snell *et al.*, 1994). Gap detection tests are well-established methods for assessing the temporal acuity of the auditory system. One model explaining the limits for gap detection (Green, 1985; Plomp and Bouman, 1959) is the gentle decay of sensory activation in the auditory periphery. The ringing of the auditory filter needs some time before returning to baseline excitation values at the end of the leading sound. Discontinuities can only be sensed, if the trailing sound begins after the excitation has decayed deeply enough. In this case the measure of gap duration is converted into the sensation of a drop in loudness and the detection of a gap in an otherwise homogeneous sound does not necessarily require explicit temporal processing. This model describes the dependency of gap detection performance on stimulus frequency appropriately.

Gap detection performance changes completely if leading and trailing sound markers, which limit the gap, are separated in disjunctive spectral bands. In this case both stimuli do not interact with each other in the same cochlear filter, the peripheral model of gap detection no longer explains the perceptual performance, and it is assumed that explicit temporal computation in the central auditory system is required for ''between channel'' gap detection. The concept of a ''channel'' is not restricted to different frequencies of the leading and trailing sound but it is extended, for instance, to the case of two sound sources in space (Grose *et al.*, 1999; Phillips *et al.*, 1998). In general the GTs are substantially longer in the ''between channel'' than in the ''within channel'' condition. The GT increases from a couple of milliseconds to about 30 ms if the leading burst is short. However, the listener benefits from a longer duration of the leading burst up to 100 ms (Phillips *et al.*, 1998). A similar temporal integration effect for gap detection performance has not been observed in the ''within channel'' condition: indeed there is evidence for a contrary effect. When stimulating with a noise burst containing a gap of 17-ms duration the auditory nerve responses to the onset of the trailing burst decreased slightly with increasing duration of the leading burst (Smith, 1977). This effect was explained using adaptation at this level of the auditory system. Subsequently, behavioral tests have demonstrated decreased detectable gap duration only for increasing duration of the leading burst of up to 5 ms but increasing gap duration for further increase of the leading burst duration (Schneider and Hamstra, 1999). Thus enhanced gap detection capability related to the duration of the leading burst seems to be specific to central gap detection.

One attempt to handle the apparent contradiction of different time-constants explaining temporal integration and temporal acuity is the introduction of a reset mechanism. White and Plack (1998) investigated temporal integration in detecting the pitch of unresolved harmonics. When increasing the duration of the complex tone from 20 to 40, 80 and 160 ms, the decrease of the detection threshold was higher for the unresolved than for the resolved harmonics. For short

[a]Electronic mail: bernhard.ross@ieee.org

durations the effect was even stronger than the threshold decrease by a factor of *n* (for an *n*-fold increase of the duration) expected from a central multiple-looks model. A significantly smaller threshold decrease was found for a pair of 20-ms bursts, separated by a gap of 5 ms, than for a tone of 40-ms duration, which was also inconsistent with the multiple-looks model. A reset mechanism for the central pitch processor was postulated by White and Plack (1998). When the complex tone is continuous, the pitch processor uses a long integration time constant. Any discontinuity on the order of 5 ms or longer causes the pitch processor to reset and use a short-duration discrete sampling approach.

Another reset of auditory integration was described by Clifton *et al.* (1989). When a series of clicks and a delayed copy of the stimulus were presented with speakers at different locations, the first pairs of clicks were heard as separated; the sounds of constant delay fused into a single object. The experiments could reset this mechanism if they altered the spatial arrangement of the stimulus presentation. Again, the first few clicks were perceived as separate objects and than a new integration process fused the sounds into a single object. Bregman *et al.* (1994) suggested that the auditory system may use neural offset responses to reset itself and carry out a new analysis.

There is little human electrophysiological data in the literature that are pertinent to these aspects of the temporal processing in the central auditory system. Most EEG or MEG studies have used a ''within-channel'' paradigm and demonstrated event related potentials or fields as a correlate of the perception of gaps. Several studies demonstrated a mismatch negativity, if the gap duration was infrequently changed within a sequence of stimuli with gaps of otherwise constant duration (Bertoli *et al.*, 2001, 2002; Desjardins *et al.*, 1999; Trainor *et al.*, 2001). Transient middle latency responses (MLRs) from primary auditory cortex, which is assumed to play an important role in auditory temporal processing, were recorded in response to gaps in noise with durations above (but close to) the behavioral threshold (Rupp *et al.*, 2002).

A gaplike stimulus was used in studies investigating the temporal dynamics of auditory steady-state responses (ASSRs) to click stimuli presented with 40-Hz repetition rate (Makeig and Galambos, 1989; Makeig *et al.*, 1992). When a single click was omitted in the series of stimuli a long lasting perturbation of the ASSR has been observed. The ASSR amplitude dropped down and recovered within about 200 ms. The ASSR phase also showed a systematic initial deviation of about 60° which returned to zero after some 100 ms. This observed effects of an omitted stimulus resembles the time-course of ASSR after the onset of an amplitude modulated (AM) tone-burst, which was studied recently in more detail (Ross *et al.*, 2002) by using AM tone-bursts of 250- and 500-ms duration presented with a stimulus onset asynchrony (SOA) of 0.5, 1.0 and 2.5 s. A transient gamma-band response and the linearly rising slope of the ASSR seen in the response signal were assumed to be distinct components with different functional meaning. The rising slope of the ASSR of 200-ms duration was interpreted as reflecting temporal integration in the central auditory system. Further results of the study (Ross *et al.*, 2002) were that the time-course of the ASSR did not change when the interstimulus interval (ISI) was shortened to 250 ms. After stimulus offset, the ASSR amplitude also returned to baseline in a much shorter time than needed for the response onset. From the similarity between the event related perturbation observed by Makeig and Galambos (1989) and the onset response observed in our previous study (Ross *et al.*, 2002) a hypothesis has been formulated that an auditory event causes a reset of the ASSR followed by an onset response.

The objective of the present MEG study was to investigate whether the temporal integration effect seen in the time-course of the ASSR onset continues across a gap in the AM stimulus. Therefore, the time-courses of 40 Hz ASSR to sequences of AM tone-bursts were investigated for a wide range of silent intervals between bursts. This experiment was performed to determine the shortest gap in AM sound that produces an observable change in the ASSR. In principal, the ASSR is an auditory evoked activity, following the temporal structure of the stimulus. Therefore, the ASSR is an appropriate component of brain activity for studying temporal processing in the auditory system. In addition, behavioral tests were carried out to measure subjective performance in detecting a short gap inserted in an AM sound. The experimentally obtained time-courses of the ASSR were compared with the results of these behavioral tests in order to address the functional implications of the ASSR for temporal processing in the central auditory system.

II. METHODS

A. Magnetoencephalographic recordings

1. Subjects

Three experimental sessions were carried out with nine subjects each. A group of 16 right-handed subjects (nine female) between the ages of 21 and 32 years who showed the highest ASSR-amplitudes in response to 40-Hz AM tones were preselected out of 32 members of the test subject's pool. Nine subjects out of this group participated in each experimental session, respectively. None of them served as a subject in all experiments. The subjects had no history of otological or neurological disorders and had normal audiological status (air conduction thresholds no more than 10 dB HL at octave frequencies between 0.25 and 8 kHz). The study was reviewed by the Ethics Commission of the Medical Faculty of the University of Münster. Informed consent was obtained from each subject after they were instructed about the nature of the study in accordance with the principles of the Declaration of Helsinki.

2. Stimulation

The stimuli were sequences of sinusoidal amplitude modulated (AM) tone-bursts separated by short silent gaps. The fully modulated tone-bursts were generated by multiplication of an sinusoidal carrier with frequency f_c and amplitude a by a shifted cosine function with the modulation frequency f_m yielding the desired signal $y(t)$,

$$y(t) = a \cdot \sin(2\pi f_c \cdot t) \cdot (1 - \cos(2\pi f_m \cdot t)). \quad (1)$$

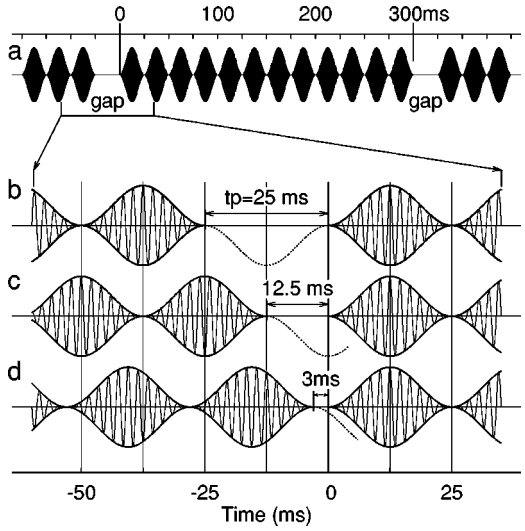


FIG. 1. Amplitude modulated stimuli. (a) The stimulus sequences consisted of bursts of 300-ms duration of 500-Hz tones fully amplitude modulated with 40-Hz modulation frequency. (b) The inserted gap of 25-ms duration is equivalent to one omitted AM period. Silent intervals of multiples of the AM period were used in the first stimulus sequence. The dashed line continues the modulation of the leading AM tone-burst and demonstrates no phase change across the gap. (c) Modulation gaps of 12.5 ms as shown here and multiples of half the modulation period were used in the second stimulus sequence. The dashed line demonstrates a phase step of 180° for odd numbered multiples of 12.5-ms gap duration. (d) The gap duration of 3 ms was the shortest value in the third stimulus sequence using gaps of multiples of 3 ms.

The duration of the AM burst was exactly a multiple of the AM period. Thus, the stimulus begun and ended smoothly at zero and no additional window function was necessary. In order to evoke maximal ASSR amplitudes the modulation frequency was selected at 40 Hz in the first three experiments investigating the influence of gap duration. For the same reason a relatively low carrier frequency of 500 Hz was used (Ross *et al.*, 2000). The AM tone-burst duration of 300 ms was equivalent to 12 periods of the amplitude modulation. With this choice the tone-bursts lasted longer than (*i*) the typical integration times of the auditory system (Gerken *et al.*, 1990), (*ii*) the onset time of the ASSR (Ross *et al.*, 2002), and (*iii*) the maximal duration of the leading sound which influences gap detection (Phillips *et al.*, 1997). On the other hand, the duration had to be made as short as possible in order to achieve a sufficient number of repeated stimulus presentations within a reasonable measurement time.

Sequences of tone-bursts interrupted by modulation gaps of different duration were used for three experimental sessions. In the first session the gap durations were multiples of 25 ms. Thus, the stimulus was equivalent to the omission of one or multiple complete periods of the 40-Hz AM sound. In the second and third sessions, the gap durations were shortened to multiples of 12.5 and 3 ms, respectively. The first two sequences consisted of nine tone-bursts and the third of 12 tone-bursts. The total length of the repeatedly presented sequences was 4 s in all conditions.

The segments of stimulus waveforms shown in Figs. 1(a)–(c) demonstrate that the gaps were inserted at the minimum of the amplitude modulation. This ensured a smooth signal envelope with no additional steps at the beginning and end of a gap. If the gap duration was an integer multiple of the AM period, the envelope time-course continued unchanged after the end of the gap. If the gap duration was equal to odd multiples of 12.5 ms a phase step of 180° occurred in the trailing AM signal compared to the envelope of the leading AM signal. This phase step became smaller with decreasing gap duration as shown in Fig. 1(c). The carrier signal of the trailing burst started with zero-phase regardless of the gap duration. The start of the onset slope after the gap defined the zero reference time for the epochs of evoked responses.

The stimuli were presented through a magnetically silent delivery system consisting of speakers (1-in. compression driver, Renkus-Heinz Inc.) mounted outside the magnetically shielded room, which were connected to a silicon earpiece through 6.3 m of echoless plastic tubing (16 mm i.d.). The transfer characteristic of this system deviated less than ± 10 dB in amplitude between 200 and 6000 Hz. Since the stimulus was defined in a narrow frequency band the stimuli were not distorted by the frequency characteristics of the sound delivery system. The transmission delay of about 19 ms was compensated by an appropriate shift of the trigger signal. Before carrying out the experimental measurements, both the signal spectrum of the stimulus and its correct timing were verified by means of a 2 cm^3 ear simulator (Brüel & Kjær model 4157) that was equipped with a $\frac{1}{2}$-in. condenser microphone (Brüel & Kjær model 4134) and connected to the silicon earpiece at the end of the sound delivery system. For the AM-tones these measurements assured correct modulation phase and modulation depth.

The stimuli were presented monaurally to the subject's right ear at a stimulus intensity of 60 dB SL (referred to the individual sensation threshold). For this purpose, the subject's hearing threshold was measured prior to each experimental session by applying the relevant stimulus tone-bursts at an ISI of 1 s through the sound delivery system.

Fifty stimulus sequences were combined into a block of 200-s duration. Between presentations of two blocks the technician conversed with the subject in order to keep her or him in an alert state. Twelve blocks were presented in the first half of an experimental session. After a break of about 10 min a repeated measurement was carried out. In total, the responses to 1200 repeated stimulus sequences were recorded in a complete session. In addition, 128 evoked responses to 500-Hz tone-bursts of 500-ms duration and 3.5-s interstimulus interval were recorded twice in each session as reference measurements. The whole procedure lasted about 2 h.

3. Data acquisition

Recordings were performed in a magnetically and acoustically shielded room. The subjects rested in right lateral position on an air mattress, with their head lying on a mould to permit stable fixation throughout the whole experimental session. The MEG was recorded with a 37-channel neuromagnetometer (MAGNES, 4D-Neuro-Imaging Inc.). The detection coils of 20-mm diameter were arranged in a circular concave array with a diameter of 144 mm and a spherical radius of 122 mm. The sensors were configured as

a first order axial gradiometer with a baseline of 50 mm. The spectral density of the intrinsic noise was about 7 fT/$\sqrt{\text{Hz}}$ above 1 Hz.

The sensor array was placed contra-laterally to the side of stimulation over the left temporal plane, centered over a point about 1.5 cm superior to the position T3 of the 10–20 system for electrode placement, as close to the subject's head as possible. During the MEG session, subjects watched cartoon videos that were projected via fiber-optic cable onto a nonmagnetic display. The subjects were instructed to stay in a relaxed state to reduce the influence of myogenic activity on the MEG-signals and their compliance was verified by video monitoring. The reason for maintaining wakefulness was that substantially decreased ASSR amplitudes during sleep have previously been reported (Jerger *et al.*, 1986; Linden *et al.*, 1985).

The magnetic field data were band-pass filtered between 0.1 and 200 Hz prior to sampling at the rate of 520 s^{-1} and 16-bit digitization and were recorded continuously.

4. Data analysis

Epochs of magnetic field data time-locked to the onset of the tone-burst following a discontinuity in amplitude modulation were selected for averaging. All epochs contained 100 ms pre- and poststimulus intervals. If the magnetic field amplitude exceeded 2.5 pT, the corresponding epoch was rejected as artifact-contaminated. In order to separate the different response components, two band-pass filters were applied to the averaged data. A 24- to 80-Hz band-pass extracted the gamma-band responses (GBRs) and the ASSR. With a second filter set between 0 and 24 Hz the low frequency P1-N1-P2 response was obtained. Zero-phase digital filtering was performed by processing the data with second order Butterworth type IIR-filters in both the forward and reverse directions. Source analysis based on the model of a single equivalent current dipole (ECD) in a spherical volume conductor was applied to the measured field distribution. Source localizations were estimated for each sampled data point in a head-based Cartesian coordinate system. The origin was set to the midpoint of the medial-lateral axis (y-axis) between the entrance of the ear canal of the left and right ear. The posterior-anterior axis (x-axis) ran between the nasion and the origin, and the inferior-superior axis (z-axis) ran through the origin perpendicular to the (x-y-plane). Estimates of the source parameters were accepted for further evaluation only if both the goodness of fit of the field of the estimated ECD to the measured magnetic field was greater than 97.5% and the distance of the ECD to the mid-sagittal plane was greater than 3 cm. The median values across all reliable source coordinates in the time interval around a magnetic field maximum and all repeated measurements were used as an estimate for the individual source coordinates. With this analysis applied to the 24- to 80-Hz band-pass filtered data the ASSR source locations were estimated. The 24-Hz low-pass filtered responses to tone-burst stimulation were used for estimation of the N1 source coordinates. Because no anatomical information was available from magnetic resonance imaging the N1 source coordinates were used as individual landmarks. The source coordinates of the ASSR were expressed as distances from these landmarks.

With the individual results of magnetic source localization spatial filters were constructed, which collapsed the time-series of the 37 MEG sensors into a single waveform of magnetic dipole moment (Ross *et al.*, 2000). This procedure was performed for each subject based on the N1-source coordinates and directions and was applied to the low-pass filtered data. Furthermore, a second filter based on the ASSR source coordinates was applied to the 24- to 80-Hz band-pass filtered response data. The polarity of the resulting dipole moment waveforms were defined in the way that the deflection around 100 ms, which is the magnetic counterpart of the slow cortical evoked N1 potential, had a negative polarity.

The time-course of amplitude $a(t)$ and phase $\varphi(t)$ of the response at the modulation frequency was obtained from the Hilbert transform applied to the 32- to 48-Hz band-pass filtered dipole moment data. The difference between response phase and stimulus envelope phase $\Delta(t)=\varphi(t)-2\pi f_m \cdot t$ indicates a time delay between stimulus and response. The mean phase difference in the steady-state time interval (250 to 300 ms after stimulus onset) was adjusted to zero. In order to reduce the effect of interindividual variation on the group statistic the individual amplitude characteristics $a(t)$ were scaled to 1.0 in the same time interval. Amplitude and phase characteristics were calculated individually and were grand averaged across the group of nine subjects. The 95% confidence limits of the grand averages were obtained from bootstrap resampling (Efron and Tibshirani, 1998).

B. Psychoacoustical tests

The GT for gaps inserted into AM sound was determined using a two-alternatives forced choice (2AFC) procedure. The stimuli were two AM bursts of 500-ms duration separated by a 900-ms silent interval. In one AM burst of the pair (randomly selected) a gap of variable duration was inserted after 250 ms. As shown in Fig. 1 the gap occurred at the minimum of the amplitude modulation in order to keep the signal envelope maximally smooth. The subject was forced to decide which AM burst contained a gap. Visual feedback was given after the subject's response. In the adaptive procedure, the gap duration was shortened after two correct answers in series and widened after a false answer. The step-size was held constant at $\sqrt{2}$. The initial gap duration of 16 ms was clearly detectable. The adaptive procedure converged to 70.7% correct answers (Levitt, 1970). The graph of the cumulative percentage of correct answers versus gap durations longer or equal to τ is an S-shaped curve $F(\tau)$ ranging from 0.707 to 1.0. $F(\tau)$ was normalized onto the range (0.0,...,1.0) and approximated by a Gaussian model $f(\tau) = 0.5\ (1+\text{erf}(\text{k}(\tau-\tau_0)))$. The parameter τ_0 is the gap duration corresponding to the 50% value of $f(\tau)$ and was chosen to be an estimate for the GT. For each test 200 stimuli were presented and about 20 min were required for a complete test. No more than three tests were performed in a session. The tests were performed using a laptop computer with audio interface and headphones (Typ AKG K141). A normative hearing threshold was determined with a group of lab members. The intensity was set 60 dB above this threshold for all

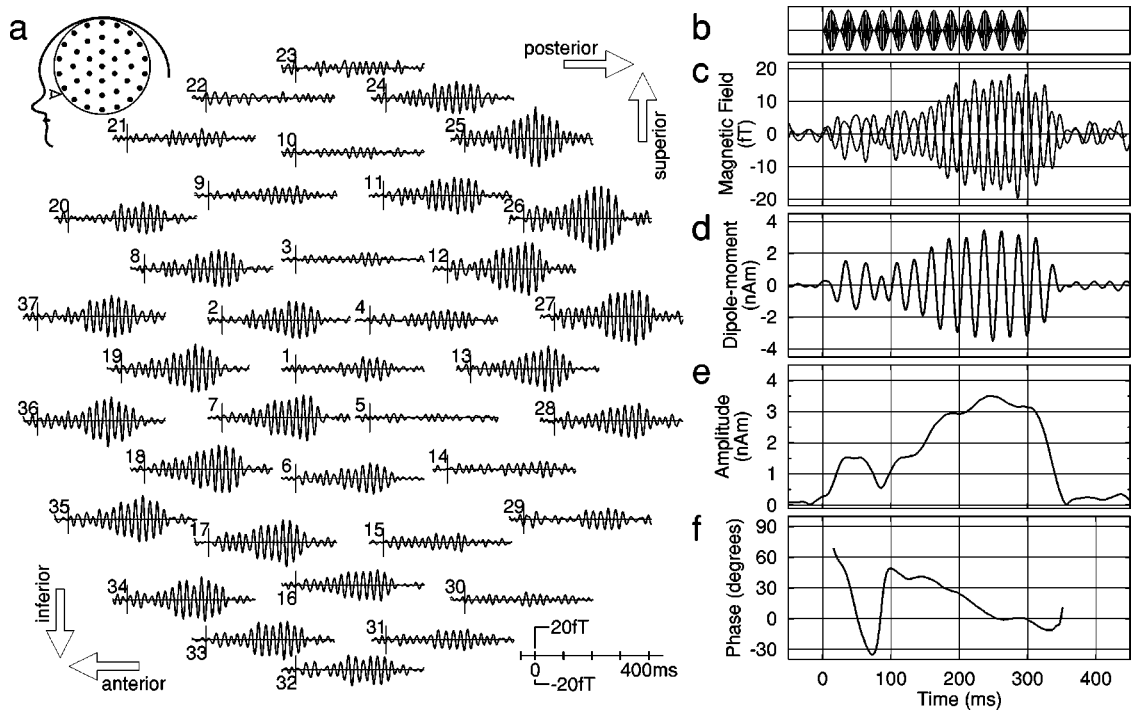


FIG. 2. Responses to AM tone-bursts filtered between 24 and 80 Hz with leading silent interval of 100-ms duration. (a) Time-series of 37-channel magnetic field data. The inset in the left top corner sketches the position of the MEG sensors over the left hemisphere. (b) Stimulus waveform, fc = 500 Hz, fm = 40 Hz, duration 300 ms. (c) Superimposed time-series of magnetic field data of channels 27 and 35, respectively. (d) Time-series of magnetic dipole moment calculated with source space projection and averaging across two repeated sessions. (e) Amplitude of dipole moments resulting from Hilbert transform applied to the time series of dipole moment shown in (d). (f) Time-series of phase difference between stimulus and response signal.

tests. In order to observe the influence of modulation frequency on the detection performance the modulation frequency was varied over the values of 16, 24, 32, 40, and 56 Hz at the constant carrier frequency of 500 Hz. In a second experiment a carrier frequency of 250, 350, 500, 1000, 2000, and 4000 Hz was used at constant modulation frequency of 40 Hz. The effect of temporal integration on the GTs was tested in the third experiment. The modulation frequency of 40 Hz and carrier frequency of 500 Hz was kept constant and the variable parameter was the time point of gap insertion after onset of the AM burst. The gap occurred at 25, 50, 75, 100, 150, 200, 250, 300, 400, or 500 ms. The burst length was extended where necessary so that the duration of the trailing sound was at least 250 ms. Six subjects participated in the first and second experiment and three in the third.

III. RESULTS

A. MEG measurements

1. Responses to single AM bursts

Clearly pronounced auditory evoked responses to the AM sound were recorded from all subjects under all stimulus conditions. A typical MEG response to a 40-Hz AM burst, which was separated by a 100-ms silent interval from the previous burst, is shown in Fig. 2. In this figure the 37 magnetic field waveforms are arranged according to the configuration of the sensor coils. The magnetic field distribution is characterized by two distinct groups of maximally responding channels with time-series of reversed polarity in the posterior-superior and the anterior-inferior region. The superimposed time-series of the MEG channels 26 and 35, which were selected out of both magnetic field maxima and are shown in Fig. 2(c), illustrate a clear polarity reversal. Such a field distribution can be sufficiently explained by a single equivalent dipole as the main source. Therefore, the localization, orientation, and dipole moment of the cortical source were estimated by approximation of a single ECD to the ASSR field distribution. The dipole localization served for source space projection through which the time-series of the cortical source activity was calculated. The source space projection allowed also the combination of response signals from two repeated measurements. The resulting waveform displayed in Fig. 2(d) exhibits four intervals with distinct temporal dynamics. The first 100 ms after stimulus onset are dominated by the transient gamma-band response. Thereafter, the ASSR amplitude developed in the interval from 100 to 200 ms. About 200 ms after stimulus onset the amplitude reached its maximum and stayed constant until the end of the stimulus. The constant response amplitude characterizes the steady state, which is terminated by the rapid decay after stimulus offset. Typical ASSR magnitudes were about 15 fT for the best responding MEG channels, and the dipole moment amplitudes were on the order of 3 nAm.

The Hilbert transform applied to the dipole moment time-series resulted in characteristics of the instantaneous amplitude [Fig. 2(e)] and phase [Fig. 2(f)], respectively. The time-course of the amplitude displays four distinct intervals: (1) a peak of the transient response around 50 ms after stimulus onset, (2) an almost linearly rising amplitude between 100 and 200 ms, (3) the interval of ASSR with only little variation of the amplitude from 200 ms to the end of the stimulus, and (4) the rapid decay of the ASSR amplitude

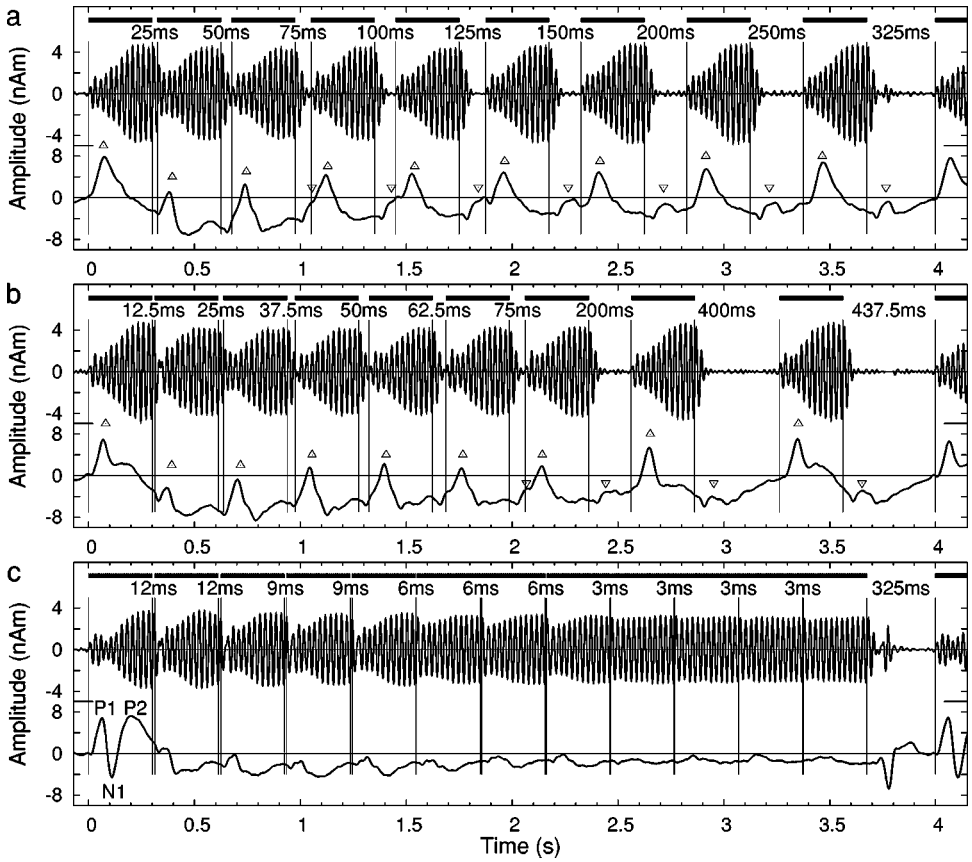


FIG. 3. Summary of grand averaged dipole moment response waveforms to sequences of AM tone-bursts. (a) Gap durations equal to multiples of the modulation period (25 ms), (b) gap durations of multiples of 12.5 ms, and (c) gap durations of multiples of 3 ms. The bars at the top of each diagram represent the time-course of AM tone-burst stimuli. The upper time-series is the 24- to 80-Hz band-pass filtered response. The lower trace shows the 0- to 24-Hz low-pass filtered response waveforms. The upward triangle symbols in (a) and (b) denote the ''on''-response in the low-pass filtered signal and the downward triangle symbols the corresponding ''off''-response. Note the twofold enlarged scale for the high frequency response compared to the low frequency response.

within less than 50 ms. The corresponding phase characteristic displays the time-course of the phase difference between stimulus and response signal referenced to its steady-state value. In the first 100 ms the phase characteristic features large variations. Although the transient response is time locked to the stimulus onset this time course indicates that the response frequency differs from the modulation frequency. Therefore, the response signal is not phase locked with the stimulus. During the interval from 100 to 250 ms after stimulus onset the response phase linearly decreases by 45°. This is equivalent to a latency increase of 3 ms during the interval of increasing ASSR amplitude. In the latency interval from 250 to 300 ms the constant phase difference between stimulus and response denotes the steady state. Thus the amplitude and the phase characteristic allows determination of the point in time at which the response signal reaches the steady state.

2. Responses to sequences of AM bursts

The response waveforms obtained from sequences of multiple AM bursts are summarized in Fig. 3. The responses to the first stimulus sequence with ISIs equal to exact multiples of a modulation period (25 ms) are shown in Fig. 3(a). Accordingly, the ISIs in the second sequence [Fig. 3(b)] were multiples of 12.5 ms and were reduced to even shorter gaps of multiples of 3 ms in the third sequence [Fig. 3(c)]. The bars at the top of the three diagrams denote the presentation intervals of the 300-ms AM burst within each periodically repeated stimulus sequence of 4-s duration. The upper traces show the 24- to 80-Hz band-pass filtered and the lower traces the 24-Hz low-pass filtered dipole moment waveforms. For ISIs longer than about 60 ms the responses to single AM bursts were clearly separated from the response to the preceding stimulus. This is the case for the six last stimuli in the sequence with gap durations of multiples of 25 ms shown in Fig. 3(a) and for the last four stimuli in the sequences with gap durations of multiples of 12.5 ms shown in Fig. 3(b). Some overlap between responses to successive stimuli occurred for shorter gaps. This became obvious from the response waveforms obtained in the third experiment when the gap durations were as short as 3 ms [Fig. 3(c)]. However, a perturbation of the ASSR amplitude can be observed even at 3-ms gap duration, which was the shortest disruption of the AM sound. The effect was more pronounced at the 6-ms gap, and the gap of 12-ms duration forced the ASSR amplitude to decay to the baseline.

Apparently, the effect of ASSR amplitude reduction after an inserted gap was smaller after the 25-ms gap, that is the second stimulus in the first sequence and the third in the second sequence, compared to the shorter 12.5-ms gap. At the 25-ms gap the leading and the trailing stimulus were in phase. Therefore, the overlapping components of the responses to both stimuli added up optimally. In contrast, at the 12.5-ms gap the leading and the trailing stimulus were of opposite phase. Consequently, the overlapping responses cancelled out each other partially.

A remarkable observation from these response signals was the almost constant recovery time of ASSR amplitude after disruption of the stimulus signal. Even after the 6-ms gap, the ASSR needed about 200 ms to recover to the final value. This was about the same time needed for the ASSR to develop after the onset of a single AM burst.

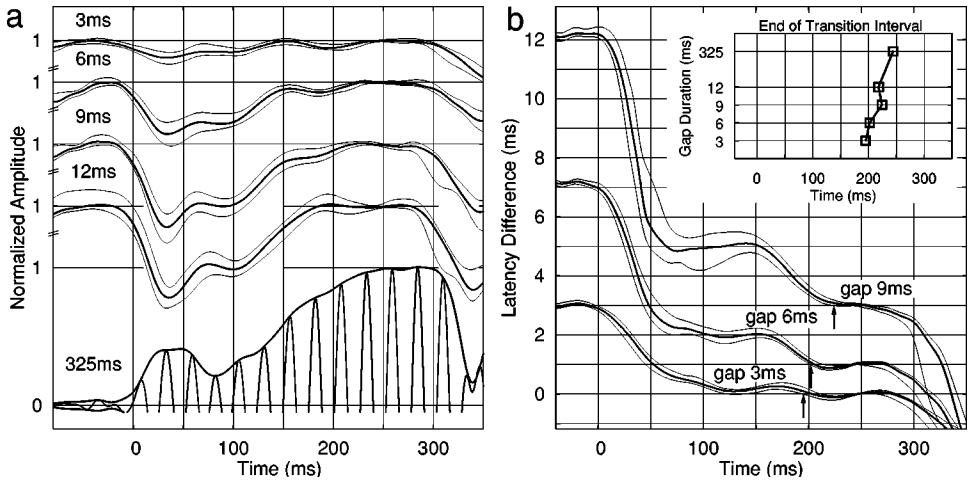


FIG. 4. (a) Group averages of time-courses of the response amplitude obtained with the third stimulus sequence. The lower trace shows the 24- to 80-Hz band-pass filtered response to the first stimulus within the sequence (ISI 325 ms) and the response signal envelope obtained by Hilbert-transform. The individual response amplitudes were scaled to the mean of one in the interval 250 to 275 ms after stimulus onset. The graphs of response signal envelopes to stimuli following gaps of 12-, 9-, 6-, and 3-ms duration, respectively, which are shown with thick lines, are shifted upward in order to avoid overlapping between the graphs. Therefore, the figure displays the deviation of the response amplitude from the mean in the steady-state interval. The 95% confidence limits for the mean across the group are shown with thin lines. (b) Time-course of the phase difference between the AM stimulus and the response signal after short duration gaps. The *y*-axis was scaled according to the apparent latency at 40 Hz (25 ms equals 360°). The grand average phase characteristics are shown with thick lines, the thin lines show the 95% confidence limits of the mean. The phase characteristic obtained after 6-ms gap was shifted by 1.0 ms and the 9-ms gap characteristic by 3.0 ms upward in order to avoid overlapping graphs. The arrows point to the intersection of the lower confidence limit and zero which defines the beginning of the steady state. The upper right inset shows these time points for the different gap duration of the third stimulus sequence.

3. Perturbation of ASSR after short interruption of AM sound

Although some overlap of succeeding responses was visible when the gap duration was 12.5 ms or longer [Figs. 3(a) and (b)] all response waveforms showed a fast decay of the ASSR amplitude. Independently of the gap duration, the ASSR amplitude recovered during a 200-ms interval after the start of the trailing AM burst. For shorter gap durations the ASSR amplitude diminished only partially as shown in Fig. 3c. The perturbation of the ASSR amplitude after short gaps was investigated in more detail in order to find the shortest gap that still elicits a detectable change in the ASSR amplitude.

Group averages of the time-course of the amplitude obtained from the Hilbert transform are shown in Fig. 4(a) together with the 95% confidence for the mean. The waveforms demonstrate a significant reduction of ASSR amplitude in the first 100 to 200 ms. The steady state is given by the time interval during which the horizontal line of 1.0 lays between the confidence limits for the mean amplitude. The first intersection of the confidence limits with the amplitude value of 1.0 defines the end of the transition into the steady state. For the various gap durations this point is reached between 150 and 250 ms after the start of the trailing AM burst. However, the amplitude characteristics show some ripple, which did not allow us to determine reliably the start of the steady state in all cases. The characteristics of the phase difference between stimulus and response signal [Fig. 4(b)], which were obtained from the same Hilbert transform, showed an unambiguously defined end to the transition interval. The obtained duration was between 195 ms in the case of a 3-ms gap and 220 ms in the case of 12-ms gap duration and increased to 245 ms for the 325-ms ISI.

The area between the amplitude characteristic and the maximum value of 1.0 was calculated as a quantitative measure for the perturbation of the ASSR amplitude in the latency interval from 100 to 200 ms. This measure is equivalent to the rectangular area of height 1.0 and width EP, which is an equivalent to the duration of complete suppression of the ASSR activity and was termed equivalent perturbation (EP). For all investigated gap durations the values of EP are summarized in Fig. 5(b). For gap durations longer than 50 ms the EP was not considerably smaller than the maximum of about 100 ms. The ASSR perturbation diminished noticeably for shorter gaps. The 3-ms gap, which was the shortest gap in the experiment, resulted in a value of about EP $=5$ ms, which is just 5% of the maximum value. The steep slope of the characteristic in Fig. 5(b) suggests no substantial EP for gap duration smaller than 3 ms, which is likely the threshold for eliciting a change in the ASSR.

4. ASSR amplitude versus ISI

The ASSR amplitude was larger after a long ISI than after a short gap between succeeding stimuli. This effect became obvious from the amplitude difference between the first and second responses in the sequence in Fig. 3(b). The ASSR amplitude after the ISI of 437.5 ms was larger than the ASSR amplitude after the ISI of 12.5 ms. In Fig. 5(a) the ASSR amplitudes in response to the 30 AM bursts (Fig. 3), which were measured as mean in the latency interval from 200 to 300 ms, are shown versus the corresponding ISI given on a logarithmic scale. The data of two repeated measurements are shown in the figure. All amplitudes were normalized to the mean across all observations and were approximated by a regression line. The slope of 0.11 per tenfold increase of the ISI was significantly different from zero [$r^2=0.49$,$F(1,58)=7.53$,$p<0.0001$].

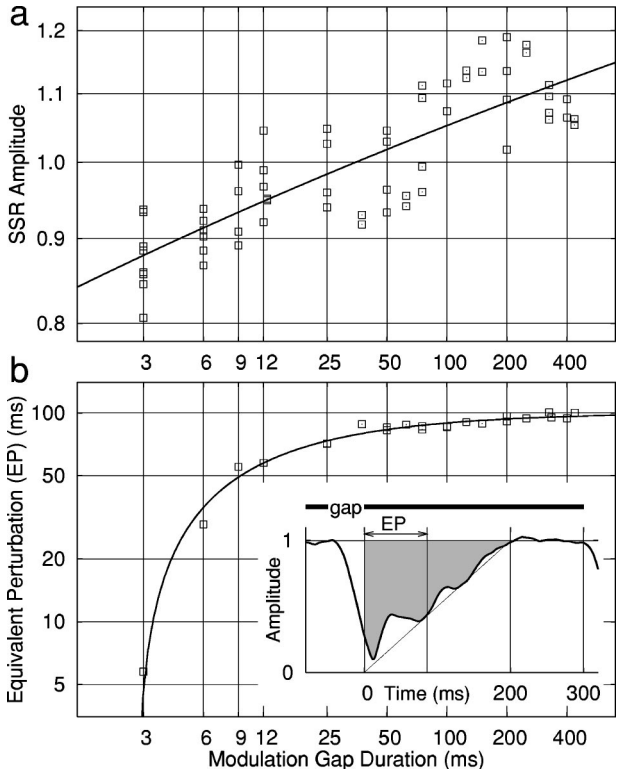


FIG. 5. (a) ASSR amplitudes as function of the duration of the preceding gap. The amplitudes are normalized to the mean of all measurements. The slope of the regression line is 0.11/decade. (b) Duration EP of complete ASSR suppression equivalent to reduction of ASSR amplitude after stimulus onset as a function of the duration of preceding gap (square symbols). The solid line is a hyperbola approximated to the data-points. The inset sketches the calculation of EP. The shaded area limited by the time interval between 0 and 200 ms and by the amplitude characteristic and the line at 1.0, denoting the normalized ASSR amplitude, is equal to the rectangle of height 1.0 and width EP.

5. A sequence of single responses

Almost no overlap between succeeding responses was seen, when the ISI was longer than 75 ms. The well separated responses were averaged to obtain a template for the mean response to a single AM burst, which is shown in Fig. 6(a). In this figure the time-course of the ASSR envelope was approximated in three intervals by straight lines representing the rising slope of 200-ms duration, the plateau, and the falling slope of 43-ms duration. Only for the first 50 ms this simplified schema poorly fits the data since the superimposed transient onset response.

Replications of the response template were combined in order to form the hypothetical responses to the stimulus sequences and are shown in Figs. 6(b)–(d) in comparison with the Hilbert-envelopes of the measured responses. For stimuli separated by gaps of $\geqslant$25-ms duration the sequence of template responses explain completely the responses to the stimulus sequence. Only for shorter gaps a shallower dip in the amplitude of the measured waveforms becomes obvious. The responses to the sequence with gaps as short as 3- to 12-ms duration demonstrate more clearly that the amplitude suppression at short gaps became smaller. Thus, for short gaps the response waveforms cannot be explained by independent responses to the leading and the trailing stimulus.

6. Low frequency response

A remarkable observation was the complete absence of an N1 wave in the low-pass filtered response signals of the stimulus sequences with ISIs of multiples of 12.5 and 25 ms, respectively. Even after the longest ISI of 437.5 ms, which is equivalent to a SOA of 737.5 ms, at the end of the second stimulus sequence only a positive response component 70 ms after onset of the trailing stimulus was the most pronounced

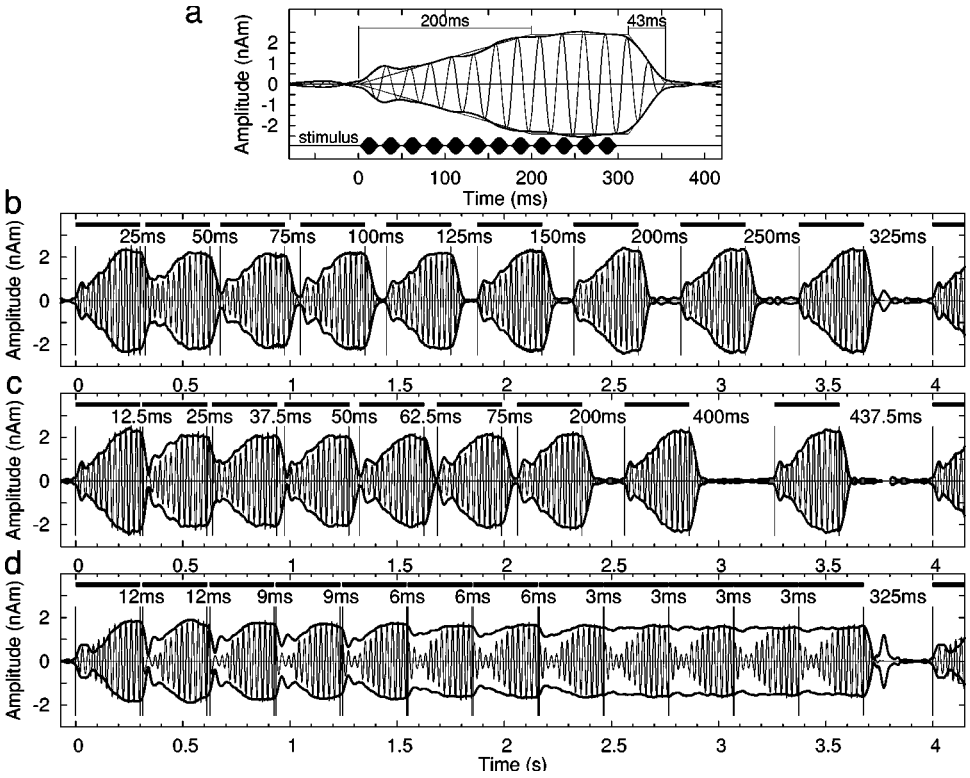


FIG. 6. (a) Response template calculated from average across all waveforms following a silent interval of more than 100 ms (thin lines). The thick lines denote the time-course of the Hilbert-envelope. The contour of rising slope, plateau, and falling slope was approximated by straight lines. The duration of the rising slope measured 200 ms and the falling slope 43 ms. (b)–(d) Synthesized responses to the sequences of AM bursts using the average response to a single burst shown in (a) (thin lines) in comparison to the envelope of the measured waveforms as shown in Fig. 3.

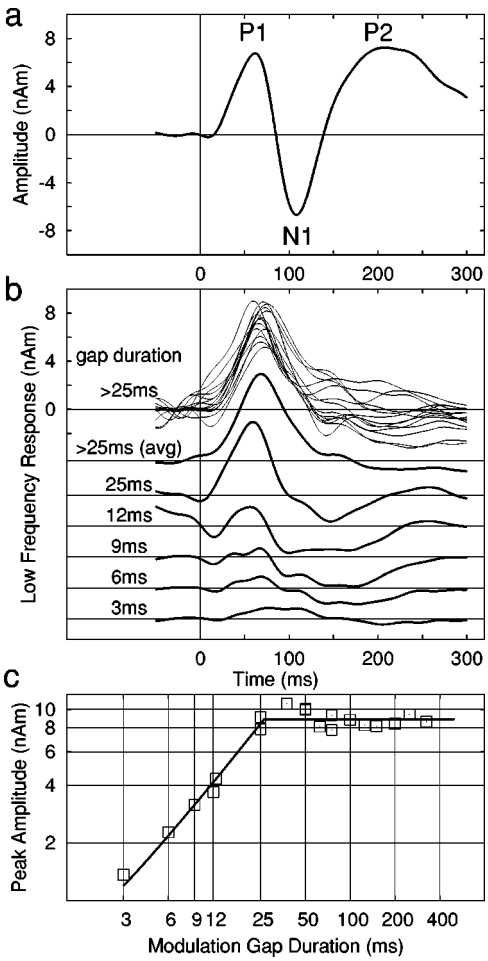


FIG. 7. The 24-Hz low-pass filtered responses to AM gaps of various durations. (a) Onset response to the beginning of the third sequence after an ISI of 325 ms. (b) Superimposed responses, which were obtained after gaps of 50 ms or longer duration in the first two stimulus sequences (top row) and average of these responses (second row). Low frequency responses to gaps of 25-, 12-, 9-, 6-, and 3-ms duration (lower 5 rows). (c) Amplitude of low-pass filtered response waveforms measured as the difference between the maximum in the interval between 50 and 100 ms and the minimum in the interval between 100 and 250 ms versus duration of the gap duration. The data points were approximated by regression lines for gap durations <25 ms and >25 ms, respectively.

wave in the response signal. This response component of positive polarity, most likely the P1-wave, was identified for all AM burst stimuli of the first two stimulus sequences and is labeled with upward pointing triangles in Figs. 3(a) and (b). Overlapping low frequency response components to successive stimuli impeded an exact measurement of the P1 amplitude. The ''off''-response, which caused such overlaps, is marked with downward pointing triangles in Figs. 3(a) and (b). The ''off''-response was almost always separated from the ''on''-response to the next stimulus for ISIs longer than 50 ms.

All low frequency responses are compiled in Fig. 7(a). Only the response to the onset of the third sequence after an ISI of 325 ms showed the typical P1-N1-P2 configuration [Fig. 7(a)]. All other response waveforms after ISIs of 50 ms or more, which are superimposed in the upper trace of Fig. 7(b), exhibited a clear pronounced P1 wave of about 8 nAm peak amplitude regardless of the actual ISI. The average across these waveforms is shown in the second row of Fig. 7(b). After a gap of 25-ms duration the P1 amplitude was still 8 nAm, however it decreased with shorter gap durations as shown in Fig. 7(b). The overall response amplitudes, which were measured as the difference between the positive peak in the time interval between 50 and 100 ms and the minimum between 100 and 250 ms, are summarized in Fig. 7(c). The amplitude characteristic was approximated by a horizontal line for gap durations longer than 25 ms. In contrast, for gap durations of 25 ms or shorter the amplitude deceased linearly, which is demonstrated by the regression line in Fig. 7(c).

Obviously, the response to the onset of the first AM burst of the third sequence shown in Fig. 3(c), which was the only one containing a N1 wave, is an exception and was excluded from the graph shown in Fig. 7(c). Although in this case the ISI of 325 ms was equal to the longest ISI within the first sequence and even shorter than two ISIs out of the second stimulus sequence, a similar N1 response was not observed in response to the first two sequences. The important difference between the stimulus sequences was that the third sequence consisted of AM bursts separated by only very short gaps. Consequently, the responses were smaller. The onset of the first AM burst in the sequence was preceded by some seconds of only incomplete responses.

7. Source localizations

Single ECD source locations for the ASSR response to 500-Hz/40-Hz AM bursts and N1 responses to 500-Hz tonebursts were estimated from 14 subjects. The mean N1 coordinates and its 95% confidence limits in the head based coordinate system were $x=6.5\pm4.4$ mm, $y=45.6\pm2.3$ mm, $z=57.2\pm4.1$ mm. In comparison to the N1 source coordinates the ASSR sources were found 1.9 ± 1.4 mm more anterior ($P=0.026$), 3.4 ± 2.3 mm more medial ($P=0.009$), and 2.1 ± 3.7 mm more inferior (n.s.).

B. Psychoacoustical tests

1. Modulation frequency

The GTs for gaps in AM sound of 500-Hz carrier frequency versus the modulation frequency in the range from 16 to 56 Hz are shown for six individuals in Fig. 8(a). The gap detection performance increased with increasing modulation frequency up to 40 Hz and was almost equal between 40 and 56 Hz. On average, the GT of 3 ms at 40 Hz corresponds to less than one-eighth of the modulation period. When expressed in proportions of the modulation period (t_p) the gap detection performance was almost constant with $0.14t_p$ at 16 Hz, $0.13t_p$ at 24 Hz, $0.12t_p$ at 32 Hz, $0.12t_p$ at 40 Hz, and $0.17t_p$ at 56 Hz. The range of variation between individual GTs was smaller than a millisecond for modulation frequencies between 32 and 56 Hz, about 3.2 ms at 24 Hz, and 4.5 ms at 16 Hz.

2. Carrier frequency

Gap detection performance with 40-Hz modulated tones is shown for six individuals in Fig. 8(b) as a function of the carrier frequency. The GT of less than 3 ms at 500 Hz corroborated the results shown for 40-Hz modulation frequency

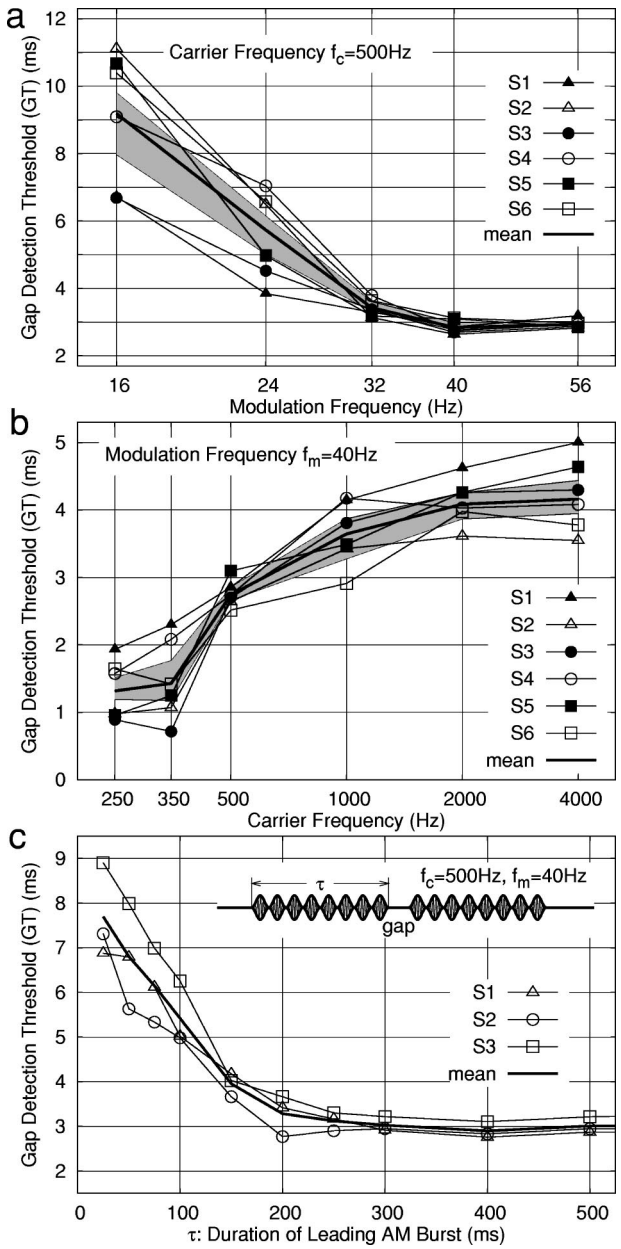


FIG. 8. (a) Gap detection performance as function of the modulation frequency of the AM stimulus for six individuals S1 to S6. The AM carrier frequency was 500 Hz. The thick line denotes the group mean and the gray shaded area the 95% confidence interval of the mean. (b) Gap detection performance as function of the carrier frequency of the AM stimulus for six individuals S1 to S6. The modulation frequency was 40 Hz. The thick line denotes the group mean and the gray shaded area the 95% confidence interval of the mean. (c) Gap detection performance versus duration of the leading AM sound as obtained from three individuals S1 to S3. The carrier frequency of the stimulus was 500 Hz and the modulation frequency was 40 Hz.

in Fig. 8(a). With increasing carrier frequency the gap detection performance decreased. At 4000 Hz the subjects were able to detect a 4.2-ms gap on the average. At 250 Hz, the lowest investigated carrier frequency, gap detection performance was highest with a GT of 1.3 ms.

3. Temporal integration

Gap detection performance increased with increasing duration of the leading AM burst. This relation is visualized with Fig. 8(c) with characteristics obtained from three individuals. If the duration of the leading burst was 250 ms or longer the gap detection performance of the three subjects reproduced the GT of 3 ms found at the 500-Hz/40-Hz AM sound and 250-ms duration of both leading and trailing AM burst. The gap duration had to be longer in order to be detected, if the trailing AM burst was shorter than 250 ms. At 25 ms, when only one period of the AM sound formed the leading stimulus, the GT was 7.8 ms, which is about 2.5 times the GT at 250 ms.

IV. DISCUSSION

The time-courses of 40-Hz ASSR to sequences of AM bursts separated by silent gaps of 3- to more than 400-ms duration were investigated. The response to a single AM burst was characterized by a 200-ms interval of developing ASSR and fast decay after the stimulus offset. Almost no interaction between the responses to succeeding AM bursts was found, if the stimuli were separated by gaps wider than 50 ms. The responses overlapped for shorter gaps; however, the perturbation of the ASSR lasted longer than the gap between stimuli. Thus, the electrophysiological threshold at 500-Hz carrier and 40-Hz modulation frequency was estimated close to 3 ms. Psycho-acoustical investigations corroborated a behavioral threshold at 3 ms for the same stimulus signal.

A. ASSR source localization

Since no structural magnetic resonance images were available from the test subjects, an overlay of the estimated source localizations with individual anatomical structures was not feasible. However, since the ASSR sources were more anterior and more medial to N1 sources, which are most likely located in lateral parts of Heschl's gyrus and the planum temporale (Godey *et al.*, 2001), the ASSR sources were assumed to be located along Heschl's gyrus, which is consistent with previous observations under the same experimental conditions (Draganova *et al.*, 2002; Engelien *et al.*, 2000; Ross *et al.*, 2002). Thus, the ASSRs obtained in this study are assumed to originate mainly from the primary auditory cortex.

B. Integrate-and-reset mechanism

The ASSR amplitude in response to an AM tone-burst stimulus was linearly rising during 200 ms after stimulus onset and decayed within less than 50 ms after the offset. Because of this asymmetry between rising and decaying slope, the underlying ASSR generator is unlikely a linear time-invariant system, for which symmetrical slopes would be expected. More probably, the ASSR generating system is time dependent and can handle different time-constants for optimal performance.

Discussions about the simultaneous use of different time-constants in auditory processing have accompanied the study of temporal properties of the auditory system for many years. Time-constants of several 100 ms were introduced to explain temporal integration, which was demonstrated with lowered hearing thresholds for sounds of increased duration (Gerken *et al.*, 1990; Green *et al.*, 1957; Plomp and Bouman,

1959). On the other hand, the models for detection of short gaps in noise use integration time-constants on the order of 3 ms (Buus and Florentine, 1985). Similar time-constants of about 2.5 ms were deduced from temporal modulation transfer functions (Bacon and Viemeister, 1985; Viemeister, 1979). Thus, the time-constants describing temporal resolution and temporal integration, respectively, are different by about two orders of magnitude. However, it seems unlikely that the brain uses multiple structures in parallel with different integration time-constants for temporal processing. This discrepancy was termed ''the integration-resolution paradox'' by deBoer (1985). Viemeister and Wakefield (1991) attempted to resolve this paradox with the model of ''multiple looks'' and hypothesized that the auditory system acquires short samples of the acoustical signal at high rates. These samples are stored in parallel and temporal integration is realized through combination of samples without the need of a long integration time constant. This attempt noticeably improved the modeling of temporal auditory processing. However, this model was not sufficient to explain all experimental observations. When measuring the discrimination performance for the fundamental frequency F_0 of an unresolved complex tone in relation to tone duration, Plack and White (2000) found higher thresholds for two 20-ms bursts, which were separated by 8 ms, than for a 40-ms burst. When the 8-ms gap was filled with noise of the same spectral range than the complex tone the two 20-ms bursts were perceived as a continuous tone and a lower threshold for F_0 discrimination was found. Plack and White (2000) discussed their results with respect to a long integration in pitch discrimination, which was reset if a discontinuity in intensity occurred.

Although previous studies described the time course of ASSR after changes in the stimulus signal, an electrophysiological correlate of the integrate-and-reset mechanism has not been described before. For instance, Makeig and Galambos (1989) reported a drop in the ASSR amplitude followed by a 200-ms recovery interval in response to a sudden discontinuity in a series of click stimuli, which were presented periodically at a rate of 40 s^{-1}. Their stimulus perturbations included abruptly changing the intensity, omitting a click, changing the ear of presentation or adding a second stimulus. Makeig and Galambos named the observed change in the ASSR the ''complex event-related potential'' (CERP) and discussed it as indicating in general the detection of a change in a continuing stimulus. The shortest gap in the study of Makeig and Galambos was the omission of one click in the series, which is equivalent to a 25-ms gap in the present study. However, in the present study the gap durations were varied over a wide range and the observed time-courses of ASSR were compared with the results of behavioral investigations. The signal-to-noise ratio of MEG data obtained in the present study was also noticeably higher than this one in the EEG and MEG data presented by Makeig *et al.* (Makeig and Galambos, 1989; Makeig *et al.*, 1992).

For gap durations equal or longer than 25 ms (one AM period) the recorded responses were reasonably explained with repeated sample waveforms, obtained from an average of responses to well separated AM burst stimuli. In contrast to the assumed long lasting integration after the stimulus onset it seems that almost no integration occurred across gaps of at least 25-ms duration. Thus, at gap duration of 25 ms and longer, the responses to the leading and trailing stimuli were entirely separated.

At shorter gap durations the separation between succeeding responses was not complete. For instance, the ASSR amplitude decreased up to 80% of its steady-state value in the case of 6-ms gap [Fig. 3(c)]. Thereafter, the time-course of the amplitude did not equal the turn from 80% to 100% of the isolated single response (Fig. 6), which would predict a recovery time of less than 50 ms. In fact the amplitude recovery required the same 200-ms period seen in an initial response. This 200-ms transition interval was even evident in the time-course of the ASSR phase at the shortest gap of 3-ms duration [Fig. 4(b)]. Thus, the ASSR perturbation induced by a short gap of less than 25-ms duration can be assumed to represent a partial reset of the ASSR amplitude followed by an onset with similar temporal characteristics as the complete onset. For gap durations longer than 25 ms the ASSR was entirely reset and the trailing burst elicited an onset response identical to that observed from wider separated stimuli. This was demonstrated in Figs. 6(b)–(d) by the successful explanation of the ''constant amplitude'' portion of the ASSR response to a sequence of AM bursts by combining the responses to a single separated AM burst.

The nonoverlapping responses following long gap durations were of similar shape. The suppression of ASSR amplitude during the rising slope was about 100 ms, which degraded to 70 ms at gap durations of 25 ms. This means that the equivalent duration of complete suppression of the ASSR was longer than the duration of the gap in the stimulus. At 6-ms gap duration the ASSR suppression was four times longer and even at a 3-ms gap the equivalent suppression of the response was twice as long. Thus the perturbation of brain responses was of substantial longer duration than the gap eliciting the response, a feature that may improve the ability to detect short gaps. The electrophysiological threshold estimated from a detectable perturbation of the ASSR waveform was around 3 ms for the stimulus with 500-Hz carrier and 40-Hz modulation frequency.

C. Low frequency response

The 24-Hz low-pass filtered responses to the stimulus sequences with gaps equal to multiples of one or a half modulation period were characterized by the almost complete absence of the N1 response, which is otherwise the most pronounced component of the auditory evoked response to stimulus onsets separated by several seconds. In the two stimulus sequences the gap duration was varied between 12.5 and 437.5 ms, which was equivalent to SOA variations between 312.5 and 737.5 ms. It is well known that N1 amplitude is substantially reduced at short SOA as compared to a SOA of several seconds. This effect is commonly explained as habituation and denotes a response of reduced strength caused by the lack of novelty, or refractoriness, which is temporarily reduced responsiveness of neural structures immediately after excitation. However, this effect is still under debate (Budd *et al.*, 1998). New interpretation is given by recent neurophysiological findings (Ulanovsky

et al., 2003) that A1 neurons in cat adapted over several seconds to repeated presentation of the same stimuli but remained hyper-responsible to deviation in frequency or intensity of the stimulus. This demonstrates that even single neurons can form time variant systems, which do not depend linearly on the stimulus parameters ISI or SOA. For periodic stimuli at SOA of 0.5 s a completely suppressed N1 was previously observed in MEG studies (Draganova *et al.*, 2002; Ross *et al.*, 2002). In these cases the P1 wave at a latency of around 70 ms became the most prominent component of the evoked response. These results were reproduced and confirmed here with the responses to the first two stimulus sequences (Fig. 3). The P1 amplitude was almost constant for gap durations of 25 ms and longer. For shorter gap durations P1 amplitude decreased and the overall amplitude of the low-pass filtered response decreased also uniformly and reached the detection threshold below the gap duration of 3 ms. This value was consistent with the electrophysiological threshold for the ASSR.

An N1 ''on''-response to the first stimulus and an N1 ''off''-response to the last stimulus within the sequence were only observed in the third experimental sequence with the shortest gap durations. The longest silent period in this sequence, before the start of the first stimulus, was not longer than in the first two sequences. Thus, the N1 response is not determined by the parameter ISI or SOA in a simple manner. Both habituation and refractoriness could explain the complete suppression of N1 responses because equal stimuli were repeatedly presented in relatively fast order. Then, the question arises, what is the novelty at the beginning of the last sequence which elicits a clearly pronounced P1-N1-P2 response? One obvious difference to the other stimulus onsets is that for several seconds the stimuli were separated by only short gaps. Consequently, only small effects were seen in the evoked responses in Fig. 3. The ASSR was only partially reset. This interval of several seconds without substantial responses might enhance the responsiveness of the N1 generating system. It is speculative and it needs further investigation whether there is a conjunction between the repeated reset of the ASSR in the experiments using the first two stimulus sequences and the complete suppression of the N1 in these sequences.

D. Gap detection thresholds

For gap detection of pure tones, Moore *et al.* (1993) reported thresholds between 6 and 8 ms for frequencies between 400 and 1000 Hz and slightly longer detectable gaps of 8–10 ms at 200 Hz and 7–10 ms at 2000 Hz (Moore *et al.*, 1993). Their thresholds depended little on stimulus intensity in the range of 40 to 85 dB SPL. In the present study the detectable gaps in AM sound were considerably shorter. Between 500 and 2000 Hz, gaps were detectable in AM sounds which had only the half of the duration of those obtained by Moore *et al.* for pure tones. In the AM sound of the present study, the GT was significantly shorter at 250 Hz than at 500 Hz, whereas in the pure tone condition of Moore *et al.* the GT increased when changing the frequency from 400 to 200 Hz. Thus, AM GT exhibited two features, which strongly contrast those for pure tones and for narrow-band noise. The first remarkable property was the increasing gap detection performance for lower stimulus frequencies. Second, GT were clearly lower for AM sounds than for pure tones. At the first glance an opposite effect might be expected, because the temporal amplitude fluctuations produced by the 40-Hz modulation are perceived as roughness, and it is likely that they interfere with the temporal structure of an inserted gap. Thus, elevated thresholds would be expected under the assumption that the same mechanism is used for AM sounds and for pure tones. Both discrepancies can be interpreted as evidence for different mechanisms of gap detection in pure tones and noise, and the detection of a discontinuity in an AM rhythm.

E. Temporal integration

The behavioral data showed a decrease in the GT, when the duration of the leading AM burst increases. For leading bursts longer than 200 ms gap detection performance remained constant. Improving gap detection performance with increasing duration of the leading stimulus restricted to a 100-ms range was previously reported for central gap processing in the ''between channel'' gap condition (Phillips *et al.*, 1998; 1997). Enhanced gap detection performance with increased duration of the leading sound has not been reported for peripheral ''within channel'' gap detection. Thus, the behavioral data provides evidence for central processing of a gap in the AM sounds used in this study, even though the stimulus did not resemble those traditionally used for the ''between channel'' condition.

Evidence for a central mechanism of gap detection in the ''within channel'' condition was provided by Eggermont (2000), who recorded single and multi-unit responses to gap stimuli from cat primary auditory cortex. The minimum detectable gap in the white noise stimuli decreased with increasing duration of the leading burst for about 100 ms. The stimulus was typical for ''within channel'' detection of a discontinuity in an otherwise homogeneous sound. However, it has been shown that the neural responses represented a property of cortical neurons. The primary auditory cortex neurons fired only when the time interval between onset of the leading and trailing burst was larger than the refractory time of these neurons on the order of about 50 ms. The interpretation of these results was that instead of the gap duration itself, which is the interval between offset of the leading to onset of the trailing sound, the time interval, which includes the leading sound and the gap, is the most critical parameter. Eggermont (2000) compared the results obtained from cat with the human behavioral data of Phillips *et al.* (1997) and concluded that central gap detection performance may be determined by the properties of primary auditory cortex neurons.

However, the case of an AM sound seems to be even more complex. Instead of a single onset response at the beginning of the leading sound, additional activity is elicited periodically by the rhythm of modulation every 25 ms.

F. AM gap detection

Studies of human gap detection performance have demonstrated widely variable measures for the GT ranging from

1 to some 10 ms. The differences between the ''peripheral'' and the ''central'' mechanisms play an important role in this variation. Gap detection also depends strongly on the experimental methods, mainly the design of the stimulus signal. It seems impossible to create a gap in a stimulus, where the gap duration is the only variable parameter with all other stimulus parameters kept constant over time. In general, the listener can use additional cues to detect the gap. Examples are changes in the stimulus spectrum, changes in the time course of the intensity, and changes in the overall signal duration.

For the stimuli used in the present study the gap was inserted at a local minimum of the AM sound. At this point the stimulus amplitude and the first derivative of the envelope are zero (Fig. 1). Thus, the transitions from the leading sound to silence and from silence to the trailing sound were smooth and the steep amplitude steps, which can cause spectral splatter, were avoided although the inserted gaps slightly change the modulation side bands of the stimulus spectrum. Yet the stimulus is still frequency specific and its spectral energy is concentrated around the carrier frequency. With an inserted gap, the stimulus intensity changes and the loudness fluctuation cue could be used for gap detection. Furthermore, gap durations different from multiples of the modulation period of 25 ms change the envelope phase. Therefore, the stimulus no longer has continuity in its periodicity (Fig. 1). Hence, a periodicity cue may be used for gap detection.

The limitation of gap detection trough ''ringing'' of the auditory filter was demonstrated as strongly pronounced phase effects on the detection of gaps in sinusoidal stimuli by Shailer and Moore (1987). In a similar manner, ''ringing'' of the modulation filter bank, proposed by Dau *et al.* (1997), may play a certain role for the amplitude modulation gap detection. Sek and Moore (2002) verified this hypothesis experimentally. However, the design of the stimulus, which should allow a study of the effect of modulation phase independently of gap duration, does not seem simple. The stimulus used by Sek and Moore in the π-phase condition exhibited a clear discontinuity in the stimulus envelope even for zero gap duration. The results could not convincingly show a modulation phase effect as seen in the experiments of Shailer and Moore (1987). Furthermore, the relatively small Q's between 1.0 and 2.0 of the filters in the model of a modulation filter-bank (Ewert and Dau, 2000) do not seem adequate to explain AM gap detection performance.

However, a clearly pronounced effect in the study of Sek and Moore (2002) was that the periodicity of the modulation is an important cue for AM gap detection. In case of a nonperiodical modulation the detectable gaps were about five times longer than in the case of periodical 40-Hz modulation.

For the stimuli used in the present study, the gaps were inserted at the minimum of amplitude modulation. This resulted in a smoother transition, when compared with the stimuli used by Sek and Moore. With this choice of modulation phase, the overall change in intensity, which may be perceived as a dip in loudness, was larger. However, the short gaps, which were detected especially at low carrier frequencies, cast doubt that the drop of intensity played an important perceptual role. The inserted gap causes additional spectral splatter; if the gap duration is not an integer multiple of the period of the carrier signal, that means a multiple of 2 ms in the case of 500-Hz carrier frequency. In such cases the gap introduces a phase shift in the carrier signal. Phase shift in the carrier occurred for example when inserting ASSR gaps of 3 or 9 ms duration. However, no discontinuities were seen in the response characteristics for the gaps of 3, 6, 9, and 12 ms duration. Furthermore, during the designing of the experiment it was tested that a carrier phase reversal at the AM trough cannot be perceived. It seems most likely that the phase shift in the amplitude modulation, which means a violation of strict periodicity, was the most important cue for the gap detection.

The GT increased with decreasing modulation frequency. However, the gap duration measured as a proportion of the modulation period was almost constant with the tendency to slightly higher performance for modulation frequencies between 30 and 40 Hz. In general, the subjects were able to detect a gap of about one-eighth of the modulation period. This result is congruent with the observations of Sek and Moore (2002) of similar modulation gap detection thresholds for $f_m = 5$, 10, and 20 Hz and an elevated threshold at 40 Hz at a carrier frequency of 4000 Hz. An interesting finding is that the gap detection depends on the context of the periodical envelope fluctuation. It seems that the absolute gap duration is not the relevant cue, but rather a proportion of the modulation period.

G. Relation between behavioral gap detection thresholds and ASSR

The results of the behavioral tests and the electrophysiological data were found to be consistent in at least three points. First, the GT for a gap in AM sound of 500-Hz carrier and 40-Hz modulation frequency was found close to 3 ms with little variation between subjects. The threshold for detection of a change in the ASSR evoked with the same AM sound was also close to 3 ms. Second, the ASSR amplitude increases toward lower carrier frequency especially for carrier frequencies below 1 kHz (Ross *et al.*, 2000, 2003). The dependency of the behavioral threshold on the carrier frequency showed the same characteristic. Furthermore, the gap detection performance increased with the duration of the leading burst for about 200 ms, which is consistent with the time-course of the rising slope of the ASSR onset. The time-course of the ASSR onset was recently discussed in relation to temporal integration in the auditory system (Ross *et al.*, 2002).

V. CONCLUSION

The coincidences between behavioral gap detection thresholds and ASSR amplitude were interpreted as functional relevance of ASSR for processing of AM sounds in the central auditory system. The temporal integration, revealed by improved gap detection performance after increased stimulus duration, pointed to central processing of the temporal sound structure. The time-course of the ASSR amplitude following a gap in the AM stimulus resembled the time-course of the ASSR amplitude after stimulus onset. The onset-like response after a gap gave evidence for a reset of

auditory activity evoked by the gap. The consideration about a ''reset and integrate'' mechanism may contribute to the discussion of the apparent contradiction between the high temporal acuity and the temporal integration in the auditory system.

ACKNOWLEDGMENTS

We thank Dr. Jyrki Mäkelä for the critical comments on an earlier version of the manuscript and Karin Berning for the help to carry out the experimental work. This work was supported by grants from the Deutsche Forschungsgemeinschaft (Nos. Pa392/7-2 and Ro2401/1-3).

Bacon, S. P., and Viemeister, N. F. (**1985**). ''Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners,'' Audiology **24**(2), 117–134.

Bertoli, S., Smurzynski, J., and Probst, R. (**2002**). ''Temporal resolution in young and elderly subjects as measured by mismatch negativity and a psychoacoustic gap detection task,'' Clin. Neurophysiol. **113**(3), 396–406.

Bertoli, S., Heimberg, S., Smurzynski, J., and Probst, R. (**2001**). ''Mismatch negativity and psychoacoustic measures of gap detection in normally hearing subjects,'' Psychophysiology **38**(2), 334–342.

Bregman, A. S., Ahad, P. A., and Kim, J. (**1994**). ''Resetting the pitch-analysis system. 2. Role of sudden onsets and offsets in the perception of individual components in a cluster of overlapping tones,'' J. Acoust. Soc. Am. **96**(5, Pt 1), 2694–2703.

Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., and Michie, P. T. (**1998**). ''Decrement of the N1 auditory event-related potential with stimulus repetition: habituation vs. refractoriness,'' Int. J. Psychophysiol **31**(1), 51–68.

Buus, S., and Florentine, M. (**1985**). ''Gap detection in normal and impaired listeners: The effect of level and frequency,'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer, Berlin), pp. 159–179.

Clifton, R. K., and Freyman, R. L. (**1989**). ''Effect of click rate and delay on breakdown of the precedence effect,'' Percept. Psychophys. **46**(2), 139–145.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997**). ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. **102**(5, Pt 1), 2906–2919.

deBoer, E. (**1985**). ''Auditory time constants: A paradox?'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer, Berlin), pp. 141–158.

Desjardins, R. N., Trainor, L. J., Hevenor, S. J., and Polak, C. P. (**1999**). ''Using mismatch negativity to measure auditory temporal resolution thresholds,'' NeuroReport **10**(10), 2079–2082.

Draganova, R., Ross, B., Borgmann, C., and Pantev, C. (**2002**). ''Auditory cortical response patterns to multiple rhythms of AM sound,'' Ear Hear. **23**(3), 254–265.

Eddins, D. A., Hall, III, J. W., and Grose, J. H. (**1992**). ''The detection of temporal gaps as a function of frequency region and absolute noise bandwidth,'' J. Acoust. Soc. Am. **91**, 1069–1077.

Efron, B., and Tibshirani, R. J. (**1998**). *An Introduction to the Bootstrap* (Chapman & Hall/CRC, Boca Raton).

Eggermont, J. J. (2000). ''Neural responses in primary auditory cortex mimic psychophysical, across-frequency-channel, gap-detection thresholds,'' J. Neurophysiol. **84**(3), 1453–1463.

Engelien, A., Schulz, M., Ross, B., Arolt, V., and Pantev, C. (**2000**). ''A combined functional in vivo measure for primary and secondary auditory cortices,'' Hear. Res. **148**(1-2), 153–160.

Ewert, S. D., and Dau, T. (**2000**). ''Characterizing frequency selectivity for envelope fluctuations,'' J. Acoust. Soc. Am. **108**(3, Pt 1), 1181–1196.

Gerken, G. M., Bhat, V. K., and Hutchison-Clutter, M. (**1990**). ''Auditory temporal integration and the power function model,'' J. Acoust. Soc. Am. **88**, 767–778.

Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P., and Liegeois-Chauvel, C. (**2001**). ''Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients,'' Clin. Neurophysiol. **112**(10), 1850–1859.

Green, D. (**1985**). ''Temporal factors in psychoacoustics,'' in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer, Berlin), pp. 122–140.

Green, D., Birdsall, T., and Tanner, W. (**1957**). ''Signal detection as a function of signal intensity and duration,'' J. Acoust. Soc. Am. **86**, 961–970.

Green, D. M., and Forrest, T. G. (**1989**). ''Temporal gaps in noise and sinusoids,'' J. Acoust. Soc. Am. **86**, 961–970.

Grose, J. H., Hall, III, J. W., and Buss, E. (**1999**). ''Modulation gap detection: effects of modulation rate, carrier separation, and mode of presentation,'' J. Acoust. Soc. Am. **106**, 946–953.

Jerger, J., Chmiel, R., Frost, Jr., J. D., and Coker, N. (**1986**). ''Effect of sleep on the auditory steady state evoked potential,'' Ear Hear. **7**(4), 240–245.

Levitt, H. (**1970**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Linden, R. D., Campbell, K. B., Hamel, G., and Picton, T. W. (**1985**). ''Human auditory steady state evoked potentials during sleep,'' Ear Hear. **6**(3), 167–174.

Makeig, S., and Galambos, R. (**1989**). ''The CERP: Event related perturbation in steady-state responses,'' in *Brain Dynamics: Progress and Perspectives*, edited by E. Basar and T. Bullock (Springer, Berlin), pp. 375–400.

Makeig, S., Pantev, C., Schwarz, B., Inlow, M., Hampson, S., and Gallen, C. (**1992**). ''The auditory complex event-related field to omitted steady-state probes,'' in *Biomagnetism: Clinical Aspects*, edited by M. Hoke, S. Erné, Y. Okada, and G. Romani (Excerpta Medica, Amsterdam), pp. 165–169.

Moore, B. C., Shailer, M. J., and Black, M. J. (**1993**). ''Dichotic interference effects in gap detection,'' J. Acoust. Soc. Am. **93**(4, Pt 1), 2130–2133.

Phillips, D. P., Hall, S. E., Harrington, I. A., and Taylor, T. L. (**1998**). '''Central' auditory gap detection: a spatial case,'' J. Acoust. Soc. Am. **103**, 2064–2068.

Phillips, D. P., Taylor, T. L., Hall, S. E., Carr, M. M., and Mossop, J. E. (**1997**). ''Detection of silent intervals between noises activating different perceptual channels: some properties of ''central'' auditory gap detection,'' J. Acoust. Soc. Am. **101**, 3694–3705.

Plack, C. J., and White, L. J. (**2000**). ''Perceived continuity and pitch perception,'' J. Acoust. Soc. Am. **108**(3, Pt 1), 1162–1169.

Plomp, R., and Bouman, A. (**1959**). ''Relation between hearing threshold and duration for tone pulses,'' J. Acoust. Soc. Am. **31**, 749–758.

Ross, B., Picton, T. W., and Pantev, C. (**2002**). ''Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field,'' Hear. Res. **165**(1-2), 68–84.

Ross, B., Draganova, R., Picton, T. W., and Pantev, C. (**2003**). ''Frequency specificity of 40-Hz auditory steady-state responses,'' Hear. Res. **186**(1-2), 57–68.

Ross, B., Borgmann, C., Draganova, R., Roberts, L. E., and Pantev, C. (**2000**). ''A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones,'' J. Acoust. Soc. Am. **108**, 679–691.

Rupp, A., Gutschalk, A., Hack, S., and Scherg, M. (**2002**). ''Temporal resolution of the human primary auditory cortex in gap detection,'' NeuroReport **13**(17), 2203–2207.

Schneider, B. A., and Hamstra, S. J. (**1999**). ''Gap detection thresholds as a function of tonal duration for younger and older listeners,'' J. Acoust. Soc. Am. **106**, 371–380.

Sek, A., and Moore, B. C. (**2002**). ''Mechanisms of modulation gap detection,'' J. Acoust. Soc. Am. **111**, 2783–2792.

Shailer, M. J., and Moore, B. C. (**1983**). ''Gap detection as a function of frequency, bandwidth, and level,'' J. Acoust. Soc. Am. **74**, 467–473.

Shailer, M. J., and Moore, B. C. (**1987**). ''Gap detection and the auditory filter: Phase effects using sinusoidal stimuli,'' J. Acoust. Soc. Am. **81**, 1110–1117.

Smith, R. L. (**1977**). ''Short-term adaptation in single auditory nerve fibers: some poststimulatory effects,'' J. Neurophysiol. **40**(5), 1098–1111.

Snell, K. B., Ison, J. R., and Frisina, D. R. (**1994**). ''The effects of signal frequency and absolute bandwidth on gap detection in noise,'' J. Acoust. Soc. Am. **96**, 1458–1464.

Trainor, L. J., Samuel, S. S., Desjardins, R. N., and Sonnadara, R. R. (**2001**). ''Measuring temporal resolution in infants using mismatch negativity,'' NeuroReport **12**(11), 2443–2448.

Ulanovsky, N., Las, L., and Nelken, I. (**2003**). ''Processing of low-probability sounds by cortical neurons,'' Nat. Neurosci. **6**(4), 391–398.

Viemeister, N. F. (**1979**). ''Temporal modulation transfer functions based upon modulation thresholds,'' J. Acoust. Soc. Am. **66**, 1364–1380.

Viemeister, N. F., and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**(2, Pt 1), 858–865.

White, L. J., and Plack, C. J. (**1998**). ''Temporal processing of the pitch of complex tones,'' J. Acoust. Soc. Am. **103**, 2051–2063.

Noise-induced hair-cell loss and total exposure energy: Analysis of a large data set[a)]

Gary W. Harding[b)] and Barbara A. Bohne
Department of Otolaryngology, Washington University School of Medicine, St. Louis, Missouri 63110

(Received 19 June 2003; accepted for publication 26 January 2004)

The relation between total noise-exposure energy, recovery time, or rest during the exposure and amount of hair-cell loss was examined in 416 chinchillas. The exposures were octave bands of noise (OBN) with a center frequency of either 4 kHz at 47–108 dB sound pressure level (SPL) for 0.5 h to 36 d, or 0.5 kHz at 65–128 dB SPL for 3.5 h to 432 d. Recovery times varied from 0 to 365 d. With both OBNs, some animals were exposed on interrupted schedules. Hair-cell loss as a function of age in nonexposed animals ($N=117$) was used to correct for sensory-cell loss due to aging. For both OBNs, the ears ($N=607$) were separated into three subsets to characterize the primary hair-cell loss from noise and the secondary post-exposure loss and to determine if rest during the exposure decreased loss. Cluster and regression analyses were performed on data from the basal and apical halves of the cochlea to determine the specific rates for these three factors. It was found that: (1) when the OBN was above a critical level, there was no relation between total energy and hair-cell loss; (2) below a critical level, there were highly significant log–linear relations between total energy and hair-cell loss, but not at rates predicted by the equal-energy hypothesis; (3) rest periods during either OBN exposure reduced hair-cell loss; more so for the 4 kHz OBN than the 0.5 kHz OBN; (4) except for the highest exposure levels, the majority of outer hair cell loss from the 4 kHz OBN occurred after the exposure had terminated, while that from the 0.5 kHz OBN occurred during the exposure; and (5) a majority of the inner hair cell loss from both OBNs occurred post-exposure.
 [DOI: 10.1121/1.1689961]

PACS numbers: 43.64.Wn, 43.64.Dw, 43.66.Ed, 43.64.Bt [BLM] Pages: 2207–2220

I. INTRODUCTION

The equal-energy hypothesis is an often accepted relation between noise exposure and noise-induced hearing loss (NIHL). The notion is that the same magnitude of NIHL is produced by a variety of noise exposures, each of which has equivalent total energy. That is, the frequency content, level, and duration of the noise exposure are unimportant as long as the total energy is the same. Commonly, a 3 dB tradeoff rule is applied to estimate this relation (e.g., Eldredge and Covell, 1958). To limit damage from any given noise frequency spectrum, the duration must be halved when the level is increased by 3 dB.

The equal-energy hypothesis can also be stated with respect to noise-induced hair-cell loss. The assumption here is that permanent threshold shift (PTS) is highly correlated with hair-cell loss. The expectation is that different exposures with equal total energy produce the same amount of hair-cell loss. However, a doubling of total energy does not necessarily produce a doubling of hair-cell loss. The outer hair cells (OHC) are the first sensory cells to show signs of noise-induced damage. As the exposure continues, the OHC loss can become substantial. Even so, the correlation of OHC loss with PTS is weak (e.g., Clark and Bohne, 1978). Exposure to noise also results in inner hair cell (IHC) loss, but it is generally secondary to OHC loss and its magnitude is considerably smaller. However, IHC loss correlates with PTS very well, particularly in cases of focal IHC loss (e.g., Nordmann *et al.*, 2000) and large losses (e.g., Harding *et al.*, 2002).

A number of studies have been conducted to test the equal-energy hypothesis, but the results have often been contradictory. Findings appeared to depend upon the animal model, noise-exposure, and functional testing paradigms as well as the methods for histopathological assessment. In addition, the results have been confounded by several other issues. In some cases, the animals were terminated immediately post-exposure when there was a substantial temporary threshold shift (TTS). It has recently been indicated that the mechanism leading to TTS is completely different from that for PTS (Nordmann *et al.*, 2000). The correlation of TTS with PTS or with total exposure energy is poor (e.g., Ward, 1973). Some studies used intense, short duration noise exposures which may have produced mechanical damage or acoustic trauma. Other studies involved moderate level exposures for longer durations which produced hair-cell loss by mechanisms other than mechanical damage. Ward *et al.* (1981) observed that the equal-energy hypothesis is not applicable in cases of acoustic trauma and introduced the ''critical level'' hypothesis. That is, if the intensity of the exposure is above a critical level, the equal-energy hypothesis does not apply.

To test the equal-energy hypothesis, some studies used relatively narrow ranges of exposure level and duration. Studies have employed several different exposure paradigms ranging from pure tones (e.g., Eldredge and Covell, 1958; Goulios and Robertson, 1983), narrow-band impact noise

[a)]This work was presented in part at the 26th meeting of the Association for Research in Otolaryngology, February 2003, Daytona Beach, FL.
[b)]Electronic mail: hardingg@wustl.edu

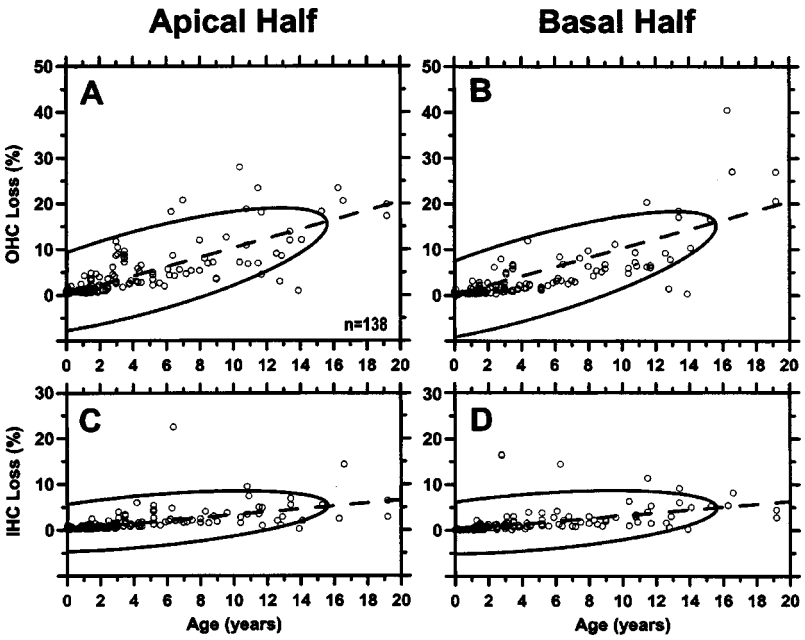


FIG. 1. Hair-cell loss due to aging for the OHC [(a) and (b)], and IHC [(c) and (d)] within the apical half [(a) and (c)] and basal half [(b) and (d)] of the OC. Regressions (q)—dashed lines; Gaussian bivariate ellipses at 0.95 probability level. Data from 117 chinchillas.

(e.g., Hamernik and Ahroon, 1998), octave bands of noise (OBN, e.g., Ahroon *et al.*, 1993; as well as impact noise), broad-band impact noise (e.g., Henderson *et al.*, 1982; 1991), and broad-band continuous noise (e.g., Salt *et al.*, 1981; as well as impact noise). Many previous studies were based upon a relatively small sample that included a large variance associated with the animals' susceptibility to noise damage. The possibility that noise affects the apical and basal halves of the organ of Corti (OC) differently was not considered. Previous studies did not distinguish between primary hair-cell loss that occurred during the exposure and secondary loss (e.g., Bohne and Harding, 2000) that continued for days post-exposure. Finally, age-related loss of hair cells was not taken into accounted.

The present study was conducted retrospectively using a large data set from the ears of chinchillas exposed to a 4 or a 0.5 kHz OBN with a wide range of levels and durations. The relation between exposure energy and noise-induced hair-cell loss was examined. The data set was large enough to determine the relevance of many of the above-noted issues and to determine the specific rates at which hair-cell loss occurred relative to total noise-exposure energy, rest during the exposure, and post-exposure recovery time.

II. DATA COLLECTION METHODS

From our collection of plastic-embedded chinchilla cochleae, the data from all non-noise-exposed control ears for animals aged from 0.2 to 19.2 yr and the data from 1 to 3 yr-old animals bilaterally exposed to noise were entered into a data set. The noise-exposed animals were subjected to an OBN with a center frequency of either 4 or 0.5 kHz in a reverberant booth on a continuous or interrupted schedule. Either immediately after the exposure or after a post-exposure recovery time, the cochleae were fixed with a buffered solution of osmium tetroxide, dehydrated, and embedded in plastic. After polymerization of the plastic, the cochleae were dissected into flat preparations from which missing hair cells were counted from apex to base and other pathology was quantified. These data were analyzed and cytocochleograms prepared relative to the percent location from the OC apex. See Bohne (1976), Clark *et al.* (1987), and Bohne *et al.* (1985; 1987; 1990) for details of noise exposures, histological processing, and quantification of hair-cell loss. The studies described here were conducted in accordance with the *ASA: Guiding principles for research involving humans or animal subjects* and were reviewed and approved by Washington University's Animal Studies Committee (Protocols #97100 and #20000131; B. A. Bohne, PI)

The information in the data set included: ear number; age (yr); OC length (mm); sound pressure level of the exposure (dB SPL); duration of exposure (total h); rest time during the exposure (total h; if any); recovery time post-exposure (d); percent IHC and OHC loss in the apical and basal halves of the OC; and an encoding of OC wipeouts, focal lesions, and nerve fiber loss in the apical and basal halves. OC wipeouts are regions of complete sensory and supporting cell loss (Bohne and Clark, 1982) and are always accompanied by nerve fiber loss that appears to be secondary to the loss of IHCs. Focal lesions are regions with at least 50% loss of OHCs, IHCs or both cell types over a distance of at least 0.03 mm (Bohne *et al.*, 1987). With substantial focal IHC loss, there is also nerve fiber loss (e.g., Bohne and Harding, 2000).

III. ANALYTICAL METHODS AND RESULTS

A. Hair-cell loss due to aging

Data from 138 non-noise-exposed ears from 117 animals, ranging in age from 0.2 to 19.2 yr, were used to determine the rate of hair-cell loss with age. Chinchillas have a life span of more than 15 yr. Figure 1 shows this relation for OHC (top) and IHC (bottom) loss in the apical [(a) and (c)] and basal [(b) and (d)] halves of the OC. Linear regressions by age were performed to determine the rates at which hair-cell loss (HCL) occurred:

$$\%\mathrm{HCL}_{\mathrm{AGE},h} = q_h x \text{ age}, \quad (1)$$

where HCL=OHC or IHC; h=OC region (basal or apical half); q=aging rate; age in years.

These rates per year were 1.010% for OHC loss and 0.346 for IHC loss in the apical half and 0.918 for OHC loss and 0.324 for IHC loss in the basal half of the OC (Pearson $r=0.862$, 0.694, 0.844, and 0.636, respectively). The apical and basal OHC losses were significantly different at the $p<0.0005$ level while the apical and basal IHC losses were not significantly different. The resulting relations were then subtracted, by OC region and hair-cell type, for the HCL data from the noise-exposed animals using Eq. (1) to remove age-related loss of hair cells. Negative values produced by this adjustment were retained to preserve the intrinsic variance. The data were plotted along with a Gaussian bivariate ellipse (derived from the eigen values of the co-variance matrix) at a probability level of 0.95. Note that the major axes of these ellipses do not correspond precisely with the regression lines. If not specified hereafter, base or basal and apex or apical refer to the basal half and the apical half, respectively. All analytical procedures were performed using SYSTAT (ver. 5.03, SYSTAT, INC, Evanston, IL). The level of statistical significance used was $p=0.05$.

B. Subsets of noise-exposed ears

The data from the noise-exposed animals were separated into subsets based upon the frequency band of the free-field exposure, the duration of exposure, the total amount of rest during an interrupted exposure, and the length of post-exposure recovery time. None of these animals were exposed at high levels for long durations. Rather, higher exposure levels were presented for shorter durations and lower levels for longer durations. For each OBN, the ears were divided into three subsets (Acute, Chronic, and Interrupted).

The Acute subset included ears exposed continuously for durations less than or equal to 9 d having a recovery time of 0 d. This subset represented the primary effects of the noise exposure (i.e., those which occurred during the exposure), minimizing overlap with secondary effects which developed with longer exposures and/or recovery times greater than 0 d. Possible secondary effects that would lead to continued degeneration in the OC include intermixing of cochlear fluids through a damaged reticular lamina (e.g., Ahmad *et al.*, 2003). With lower level, longer duration exposures, substantial secondary effects generally appear from 7 to 9 d after exposure onset.

The Chronic subset included ears exposed continuously for longer than 9 d (with any recovery time) and those ears with recovery times longer than 0 d. This subset represented the primary effects of the noise plus the secondary effects that either overlapped with the latter part of the exposure or occurred post-exposure.

The Interrupted subset included discontinuous exposures (6 h per d, 6 h per 2 d, or 6 h per week) and all recovery times. All ears in the Interrupted subset were exposed for 9 or more d. The primary and secondary effects of the exposure, rest, and recovery were expected to overlap. This subset represented the primary effects of the noise, minus the effects of rest, plus the secondary effects of recovery times greater than 0 d.

C. Noise exposures

The ears in each subset were exposed as shown in Tables I(a) and (b). Different symbols were assigned to each exposure level (or level range). Also listed are the exposure duration and post-exposure recovery time ranges and the number of ears and animals for each subset and level. The exposure levels for the 4 kHz OBN were generally lower than for the 0.5 kHz OBN because the chinchilla ear canal resonates at about 4 kHz, adding approximately 15 dB SPL at the eardrum and across the entire 4 kHz OBN (von Bismarck, 1967). For each OBN, the distribution of animals exposed at the different level and duration combinations was similar for the Acute and Chronic subsets.

For each ear, the total energy (E) to which it was exposed was calculated as follows: the exposure level was converted to pressure in pascals (Pa); the pressure was squared and multiplied by the exposure duration (excluding rest) in seconds resulting in the unit pascal squared seconds ($\mathrm{Pa}^2\,\mathrm{s}$). This unit is a common way to express exposure energy (e.g., Annex F in ANSI Standard S3.44, 1996). For example, the total energy in a 90 dB SPL, 8 h exposure is 11 520 $\mathrm{Pa}^2\,\mathrm{s}$. The log to the base 2 of the energy value was used as the independent variable [e.g., 13.5 for the above 11 520 (or $2^{13.5}$) $\mathrm{Pa}^2\,\mathrm{s}$ example]. The data were plotted on a linear X axis as powers of 2 to represent progressive doubling of total exposure energy. An example is shown in Fig. 2(a) for the noise-induced OHC loss in the basal half of the OC in the 4 kHz OBN Acute subset. For the 4 kHz OBN, exposure level was not corrected for ear-canal resonance so that the data could be plotted relative to total energy in the free field. A 15 dB SPL correction for resonance would require a multiplicative factor of $2^5\,\mathrm{Pa}^2$, which would increase the value of $\log_2(E)$ by five units.

The data in Fig. 2(a) suggest that the results could be fitted with a logistic model. However, such a model would assume that hair-cell loss has a single mechanism regardless of total energy. Because there is evidence that hair-cell loss from noise arises from more than one mechanism (e.g., Bohne and Harding, 2000), a different approach was used.

D. Cluster analysis

Inspection of a raw scatter plot (not shown) of the data for basal OHC loss from the 4 kHz OBN Acute subset clearly showed that there were three distinct groups. However, with the 108 dB SPL exposures, ears fell within each of these groups. Therefore, rather than assigning ears to a particular group based on exposure level, a cluster analysis was performed upon each subset using the K-means method (Hartigan, 1979) to separate the responses to noise, rest, and recovery into three clusters. With this method, the number of clusters must be specified beforehand. For animals having data from both ears where the two ears were initially assigned to different clusters, the ear with the lower cluster assignment was reassigned to the higher cluster. This was

TABLE I. Subsets by exposure band and exposure-level ranges, (a) 4 kHz OBN, (b) 0.5 kHz OBN.[a]

		Acute subset		Chronic subset		Interrupted subset	
	Exposure levels symbol	Exposure recovery	*N*: Ears (animals)	Exposure recovery	*N*: Ears (animals)	Exposure recovery (d)	*N*: Ears (animals)
(a)	47–72 dB SPL Pentagon	2–9 d 0 d	15 (8)	2–9 d 30 d	7 (5)	…	…
	80 dB SPL Diamond	2–9 d 0 d	10 (5)	1–9 d 6–30 d	12 (10)	36–252 0 or 30	24 (12)
	85–86 dB SPL Square	1–9 d 0 d	20 (13)	1–36 d 4–30 d	29 (22)	9–72 0–365	20 (16)
	108 dB SPL Circle	0.5–24 h 0 d	45 (33)	1–24 h 1–365 d	137 (91)	9[b] 365	2 (2)
	Total: Ears (animals)		90 (59)		185 (128)		46 (30)
(b)	65–85 dB SPL Pentagon and diamond	2–9 d 0 d	18 (9)	2–9 d 30 d	11 (8)	…	…
	95 dB SPL Square	2–9 d 0 d	27 (22)	2–433 d 7–913 d	103 (74)	36–252 0–105	79 (55)
	108 dB SPL Circle	2–9 d 0 d	4 (2)	9 d 30 d	2 (1)	9 or 36[c] 0 or 30	8 (4)
	120–128 dB SPL Triangle	3.5–14 h 0 d	7 (4)	3.5–13 h 30–730 d	27 (20)	…	…
	Total: Ears (animals)		56 (37)		143 (103)		87 (59)

[a]Grand total of 607 noise-exposed ears from 416 animals.
[b]At 92 dB SPL.
[c]At 101 dB SPL.

done so that susceptibility to noise was relative to individual animals rather than ears. After reassignment, ears ($N=15$ out of 622) that were more than four standard deviations away from the center of the cluster were treated as outliers and discarded. These latter instances showed clear evidence that the animal's two ears did not get the same exposure. It is likely that this problem arose when there was an undetected temporary or permanent conductive hearing loss in one ear but not the other.

The clusters were plotted with a Gaussian bivariate ellipse at a 0.95 probability level to enclose the majority of points within the cluster and indicate the degree of cluster overlap. Animals with two ears in the subset contributed more weight than animals with just one. However, it has been shown that there is a high correlation of hair-cell loss between bilaterally exposed chinchilla ears (Bohne *et al.*, 1986).

Figure 2(a) shows that there is no relation between OHC loss and total exposure energy for the 108 dB SPL exposure level (circles). These data form clusters 2 (blue) and 3 (red) with cluster 3 near the upper limit imposed by the ceiling effect. Cluster 3 appears to be correlated with total energy, but a regression on these data implies that a 1 Pa^2 s exposure would produce an unrealistic 70% OHC loss. Therefore, 108 dB SPL was considered to be above critical level. With exposures below 108 dB SPL, there was a clear relation between OHC loss and total energy in cluster 1 (green), which included a few noise-resistant animals exposed at 108 dB SPL. The occurrence of 108 dB SPL ears in cluster 1 required broadening the definition of critical level to include an additional mechanism for hair-cell loss; one which did not involve mechanical damage.

E. Model of noise-induced hair-cell loss

In cluster 1, the primary effect of the noise exposure on hair-cell loss was considered to be related to the total energy of the exposure. For 4 kHz OBN exposures, the most affected hair cells were the OHCs in the basal half of the OC. For 0.5 kHz OBN exposures, the most affected were the OHCs in the apical half (e.g., Bohne and Harding, 2000). The slope of the relation between hair-cell loss and total energy should be decreased by the amount of rest (R) during the exposure and increased by the amount of loss which occurred during post-exposure recovery time (P). However, there was a critical exposure level above which the relation did not apply. The following model was formed to accommodate these observations:

$$\%\mathrm{HCL}_h=[a_h \log_2(E)+k_h]+b_h \log_2(P)+c_h \log_2(R), \tag{2}$$

where HCL=IHC or OHC loss in the OC depending upon the OBN of the exposure, with the values of a, k, b, and c being determined by log–linear regression of HCL with the

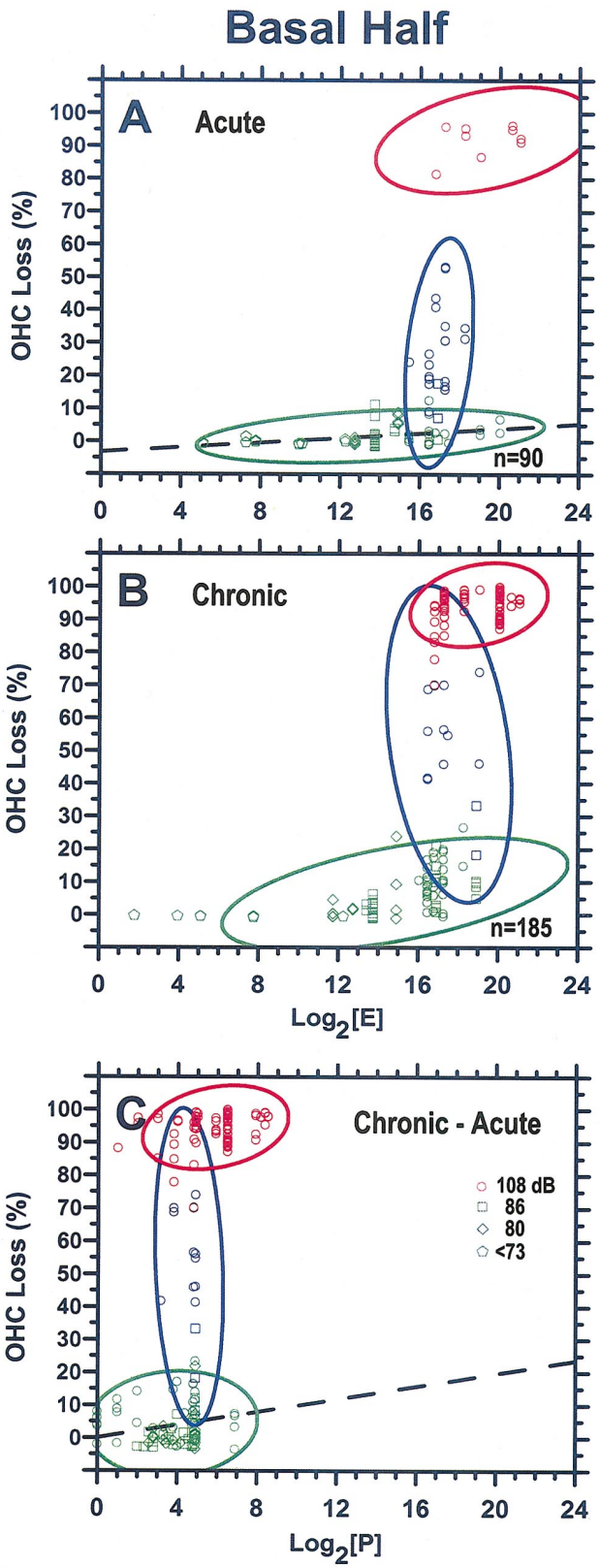


FIG. 2. 4 kHz OBN: (a) Acute subset: Cluster 1 (green) shows a strong relation (regression A) with total energy [*E*]. The exposures for clusters 2 (blue) and 3 (red) were above critical level and show no relation to *E*. (b) Chronic subset: (c) Subtraction of regression A from B for cluster 1 shows effect of post-exposure recovery time [*P*]. Symbol legend in (c) specifies exposure levels [see Table I(a)] and applies to (a)–(c). Gaussian bivariate ellipses at 0.95 probability level.

$\log_2$ of the independent variable from the data in cluster 1; h = OC region (basal or apical half).

The values of coefficient a and constant k were determined by regression [e.g., Fig. 2(a)] on the data for each of the 4 and 0.5 kHz OBN Acute subsets because E was known and both R and P were zero. Having determined a and k, the value for coefficient b was determined by regression [e.g., Fig. 2(c)] on the data for the Chronic subsets [Fig. 2(b)], cluster 1, after subtracting the Acute effect [Fig. 2(a)] using Eq. (2), with coefficients b and c equal to zero. Having determined the values for a, k, and b, the value of coefficient c was determined by regression (not shown) on the data in cluster 1 for the Interrupted subsets after subtracting the Acute and Chronic effects with Eq. (2). In the subtractions, any resulting negative values were retained to preserve the variance.

The data in Figs. 3–8 were plotted using the coefficients and independent variables on the right-hand side of Eq. (2) (without constant k) to determine the value of the independent (X axis) value. This maneuver was done to accommodate all three ongoing processes on the same scale simultaneously. For the Acute subsets, the value of coefficient a was as determined and coefficients b and c were zero. For the Chronic subsets, the values of coefficients a and b were as determined and coefficient c was zero. The P term added to the E term pushed the HCL data to the right relative to E by incorporating the duration of recovery time; moving the data slightly for shorter recovery times, more for longer recovery times. This procedure adjusted for the secondary HCL which occurred post-exposure. The relative contributions of the E and P terms can be appreciated by comparing Fig. 2(b) with Fig. 2(c). For the Interrupted subsets, the values of coefficients a, b, and c were as determined. In addition to the P term as above, the added term R pushed the HCL data to the left (because coefficient c should be negative) relative to E and P by incorporating the duration of rest during the exposure. The displacement was slight for less rest and more so for more rest. Note that plotted this way, the slopes of the major axes of the ellipses in Figs. 3–8 were changed.

F. The 4 kHz OBN

The results for the 4 kHz OBN Acute, Chronic, and Interrupted subsets are shown in Figs. 3, 4, and 5, respectively, and in Table II. The symbols in the figures represent exposure level as listed in Table I(a). Table II lists the number of ears and animals for each cluster of each subset along with the means and standard deviations for the independent variable [from the right-hand side of Eq. (2)], and the percent HCL for the basal OHC and IHC and apical OHC and IHC. Independent-, dependent-variable pairs of means are the locations of the centers for the ellipses enclosing each cluster.

Figure 3(a) shows the OHC loss in the Acute subset on an Eq. (2) scale for the basal half of the OC after cluster reassignment and outlier deletion. There were three distinct clusters as in Fig. 2(a). Cluster 1 showed a clear relation with total exposure energy. Coefficient a in Eq. (2) was 0.345 and constant k was -3.046 (Pearson $r = 0.366$). The effect of the scaling produced by coefficient a can be appreciated by com-

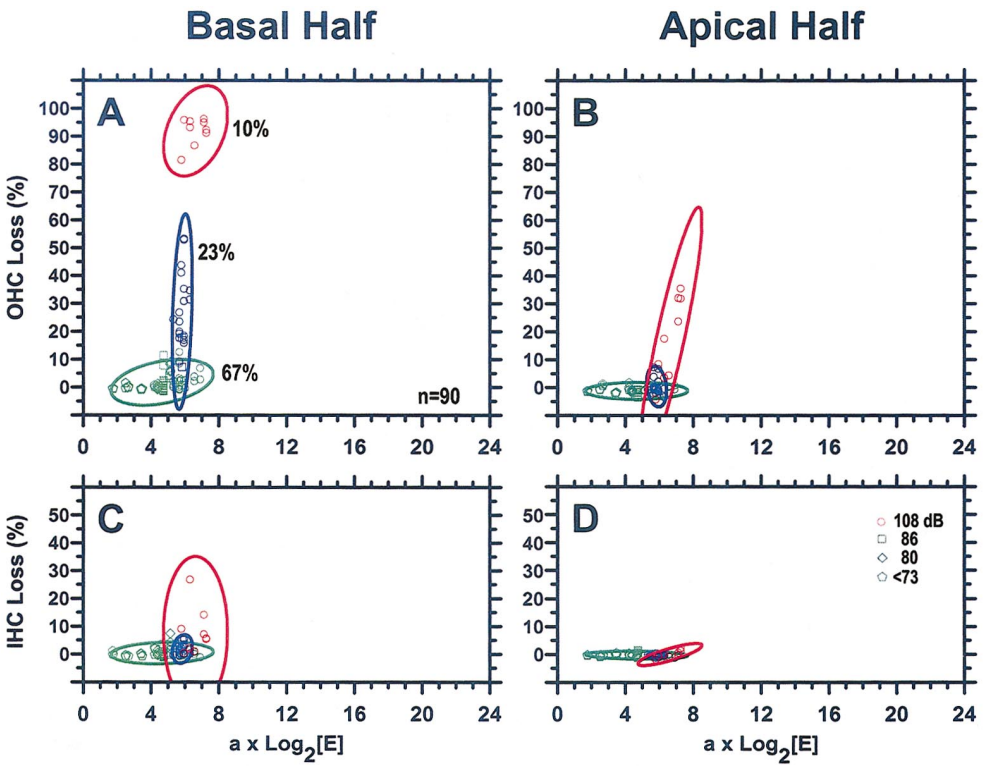


FIG. 3. 4 kHz OBN, Acute: (a) OHC—base: Below 108 dB SPL, there is a strong relation with total energy (green); 108 dB SPL (circles) is above critical level and shows no relation. (b) OHC—apex: Little OHC loss except that which spread up from base at 108 dB SPL. (c) IHC—base: Little IHC loss except at 108 dB SPL. (d) IHC—apex: Little IHC loss. Abscissa value [(a)–(d), here and in Figs. 4 and 5] from Eq. (2), basal OHC in (a). Without +15 dB SPL for ear-canal resonance. Symbol legend in (d) specifies exposure levels [see Table I(a)] and applies to (a)–(d) here and in Figs. 4 and 5. Gaussian bivariate ellipses at 0.95 probability level. Data from 59 chinchillas.

paring Fig. 3(a) with Fig. 2(a). The slope of the relation (a) was much less than would be predicted by the equal-energy hypothesis and the negative intercept (k) indicated that hair-cell loss did not begin until total energy had reached a certain magnitude. At the 108 dB SPL level (circles), clusters 2 and 3 did not show a relation with total energy. Clusters 2 and 3 had a total energy greater than $2^{15}\,\mathrm{Pa}^2\,\mathrm{s}$. Figures 3(b), (c), and (d) show the OHC loss in the apical OC and the IHC loss in the basal and apical halves, respectively. OHC loss in the basal OC spread somewhat into the apical half at the 108 dB SPL level [Fig. 3(b)]. However, there was little OHC loss at lower exposure levels [Fig. 3(b)] and even less IHC loss throughout the OC except for some at the highest exposure level in the basal half [Figs. 3(c) and (d)].

For the 4 kHz OBN exposures, clusters for IHC loss in the basal half [Fig. 3(c)] and OHC loss and IHC loss in the apical half [Figs. 3(b) and (d)] depended upon the cluster assignment for OHC loss in the basal half [Fig. 3(a)]. This dependency was based upon the observation that 4 kHz OBN exposures primarily produce OHC loss (often focal) in the basal half. Thus, the position of the other points and ellipses was relative to OHC loss in the basal half. An independent cluster analysis for the latter hair cell types and regions would have made their clusters smaller and tighter. However, the relative connections of the individual ears to the dominant region of OHC loss and the variance therein would have been lost.

Figure 4 shows the data for the Chronic subset (incorporating post-exposure recovery time) with the same arrangement as in Fig. 3. Clearly, OHC loss in the basal OC substantially increased post-exposure [Fig. 4(a)], as indicated by the changes in the proportion of the sample in the three clusters. Cluster 3 contained 51% of the sample in the Chronic subset whereas it only contained 10% in the Acute subset. The slope of the relation for cluster 1 between OHC loss and post-exposure recovery time [coefficient b in Eq. (2)] was 0.976 (Pearson $r=0.532$). The effect of the scaling produced by coefficients a and b can be appreciated by comparing Fig. 4(a) with Fig. 2(b). For cluster 1, OHC loss in the apical half [Fig. 4(b)] and IHC loss in both apical and basal halves [Figs. 4(c) and (d)] showed little or no increase. IHC loss at the highest exposure level in the basal half for clusters 2 and 3 increased dramatically [Fig. 4(c)], as did the OHC and IHC loss at the highest exposure level in the apical half [Figs. 4(b) and (d)]. Thus, the majority of the hair-cell loss from the 4 kHz OBN occurred post-exposure.

Figure 5 shows the data for the OHC loss in the basal [Fig. 5(a)] and apical [Fig. 5(b)] halves of the OC for the Interrupted subset. Because no ears were exposed to interrupted noise at a high SPL, the data were treated as a single cluster. Thus, there was no information about the effect of rest on clusters 2 and 3. In the Interrupted subset, IHC loss [Figs. 5(c) and (d)] was minimal throughout the OC. The coefficient for rest (c) in Eq. (2) was -0.923 (Pearson $r=0.772$). The negative slope indicated that rest during the exposure was beneficial.

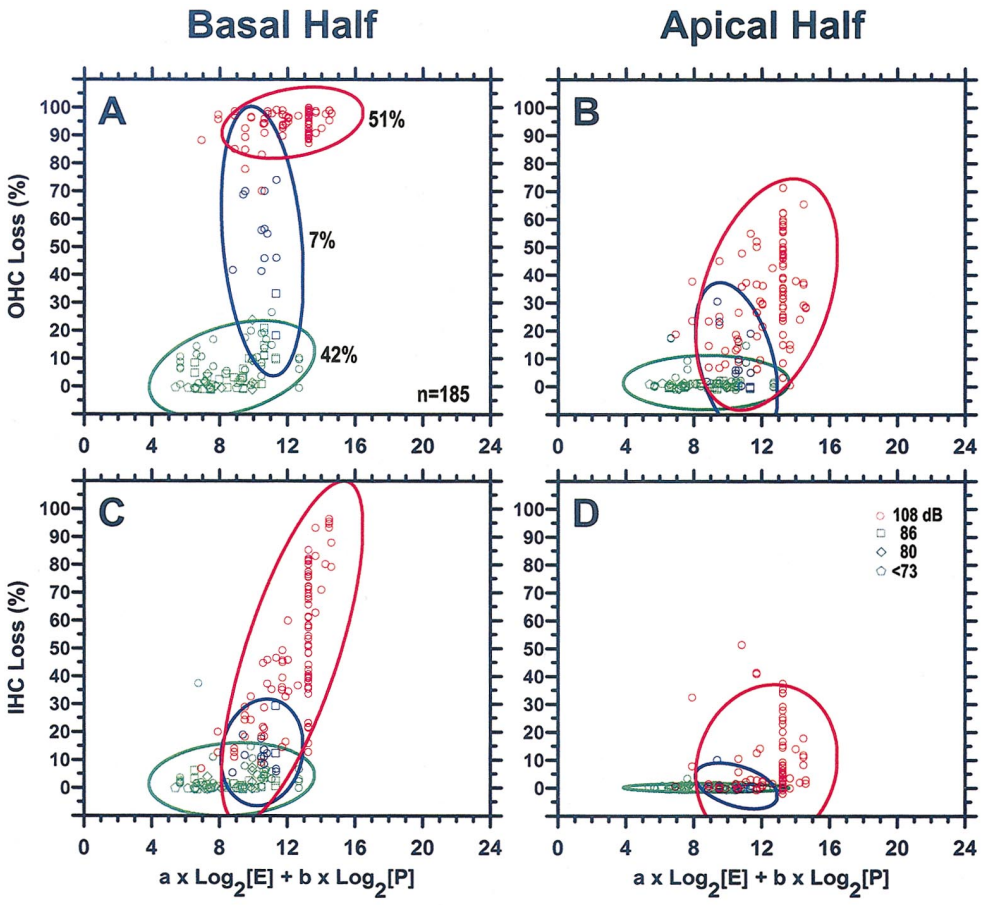


FIG. 4. 4 kHz OBN, chronic: (a) OHC—base: Strong relation with post-exposure recovery time; much more OHC loss with 108 dB SPL (circles). (b) OHC—apex: No increase in OHC loss (green) below 108 dB SPL; substantially more OHC loss with 108 dB SPL exposures. (c) IHC—base: Pattern much like OHC loss in apex. (d) IHC—apex: Pattern much like OHC loss in apex, but with much less IHC loss overall. Data from 128 chinchillas.

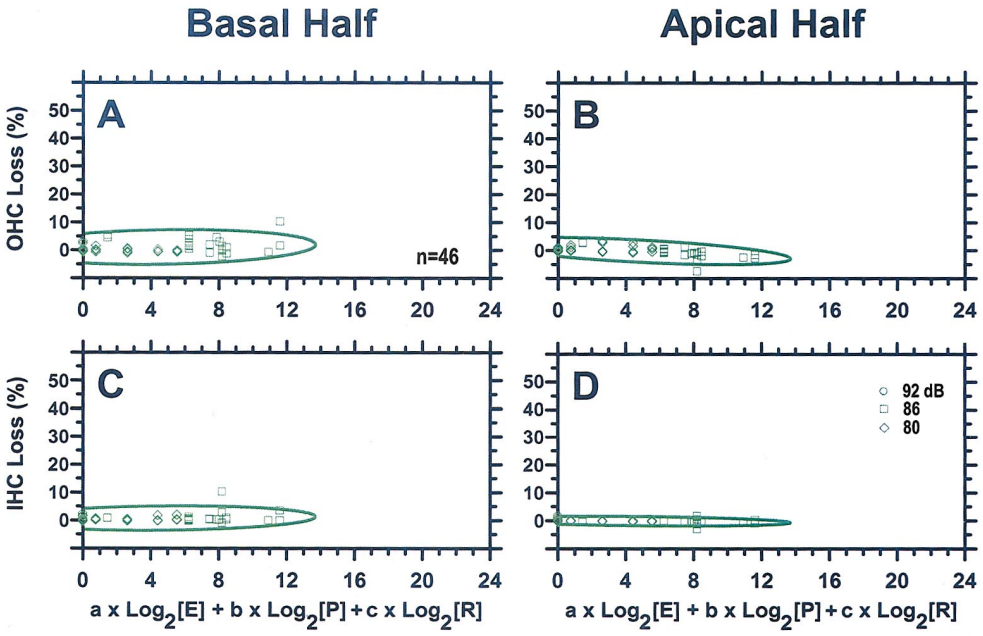


FIG. 5. 4 kHz OBN, Interrupted: (a)–(d) Hair-cell loss—base and apex: Minimal loss; in apex, somewhat less loss with longer rest periods. Exposures above 92 dB SPL not tested. Data from 30 chinchillas.

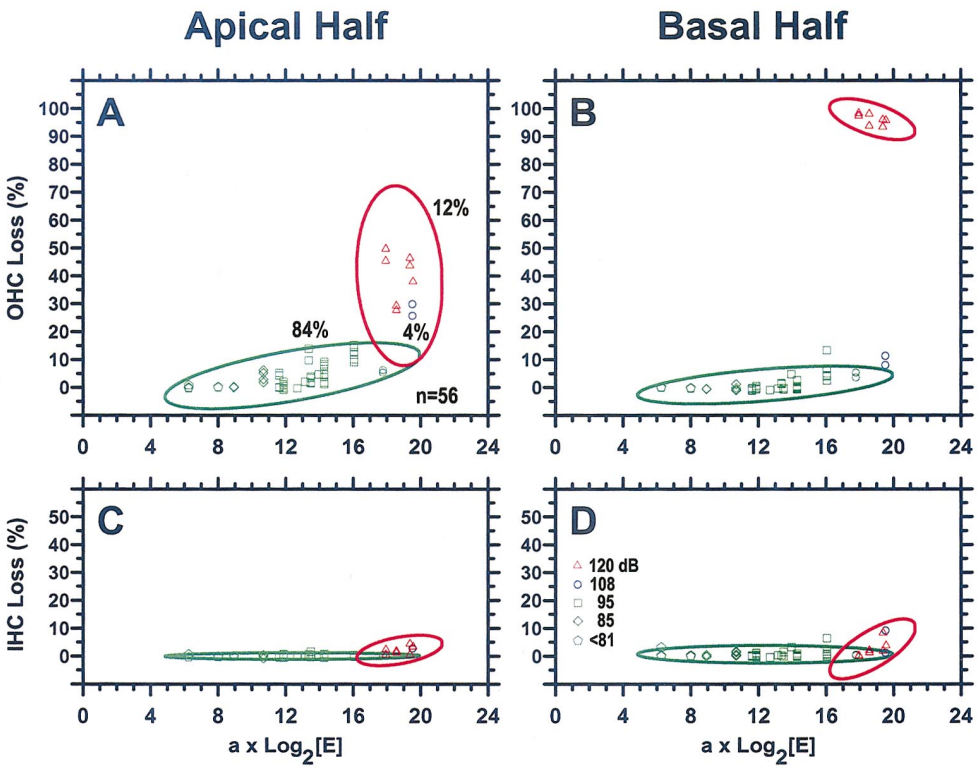


FIG. 6. 0.5 kHz OBN, Acute: (a) OHC—apex: Below 120 dB SPL, strong relation with E; 120 dB SPL is above critical level and shows no relation to E. Cluster 2 (blue) minimal and merged with cluster 3 (red). (b) OHC—base: Pattern much like basal OHC loss from the 4 kHz OBN, but at much higher E; cluster 2 merged with cluster 1 (green). (c) and (d) IHC—apex and base: Little IHC loss. Abscissa value [(a)–(d), here and in Figs. 7 and 8] from Eq. (2), apical OHC in (a). Symbol legend in (d) specifies exposure levels [see Table I(b)] and applies to (a)–(d) here and in Figs. 7 and 8. Gaussian bivariate ellipses at 0.95 probability level; ellipses for cluster 2 (blue) could not be drawn because the co-variance matrix was singular. Data from 37 chinchillas.

G. The 0.5 kHz OBN

The results for the 0.5 kHz OBN Acute, Chronic, and Interrupted subsets are shown in Figs. 6, 7, and 8, respectively, and in Table III. The symbols are as listed in Table I(b). Table III lists the number of ears and animals for each cluster of each subset, along with the means and standard deviations for the right-hand side of Eq. (2), and percent HCL for the apical OHC, IHC and basal OHC, IHC. The data are arranged as in Table II.

The OHC loss in the Acute subset for the apical half of the OC [Fig. 6(a)] showed a somewhat different pattern from that for OHC loss in the base with the 4 kHz OBN exposure [Fig. 3(a)]. There was a relation between OHC loss and total energy for cluster 1 but not clusters 2 and 3. Therefore, 120 dB SPL (triangles) was considered to be above critical level for this exposure. In cluster 1 (green), the slope of the relation for total energy in Eq. (2) was substantially larger ($a = 0.806$, $k = -8.302$, Pearson $r = 0.638$) than that for the 4 kHz OBN. In addition, cluster 2 (blue circles) was essentially missing and cluster 3 (red) appeared about where cluster 2 was with the 4 kHz OBN. Clusters 2 and 3 had a total energy greater than 2^{22} Pa^2 s. IHC loss was minimal throughout the OC [Figs. 6(c) and (d)]. The pattern of OHC loss in the basal half of the OC [Fig. 6(b)] was virtually indistinguishable from that seen with the 4 kHz OBN exposure [Fig. 3(a)], except that cluster 2 was minimally represented and more energy was required.

For 0.5 kHz OBN exposures, clusters for IHC loss in the apical half [Fig. 6(c)] and OHC and IHC loss in the basal half [Figs. 6(b) and 6(d)] depended upon the cluster assignment for OHC loss in the apical half [Fig. 6(a)]. This dependency was based upon the observation that 0.5 kHz OBN exposures primarily produce OHC loss (usually scattered) in the apical half. Thus, the position of the other points and ellipses was relative to OHC loss in the apical half.

In the Chronic subset illustrated in Fig. 7, the coefficient for recovery time in Eq. (2), cluster 1, was low ($b = 0.181$, Pearson $r = 0.173$). Thus, the vast majority of OHC loss in cluster 1 occurred during the exposure. OHC loss in cluster 3 increased post-exposure and cluster 2 appeared [Fig. 7(a)]. This secondary loss of hair cells included exposures at 95 dB SPL and higher. The relative contributions of the E and P terms was dominated by the E term and the adjustment from the P term was small. OHC loss in the basal half of the OC [Fig. 7(b)] showed a similar pattern to that from the 4 kHz OBN [Fig. 4(a)]. Much IHC loss appeared in both the apical [Fig. 7(c)] and basal [Fig. 7(d)] halves of the OC for clusters 2 and 3. However, there was little additional IHC loss in cluster 1.

The hair-cell loss for the Interrupted subset is shown in Fig. 8. Exposure levels greater than 101 dB SPL were not used, so the data were treated as two rather than three clusters. The coefficient for rest in Eq. (2) for cluster 1 was negative ($c = -0.435$, Pearson $r = 0.710$), again showing the benefit of rest during the exposure, albeit at about half the rate for the 4 kHz OBN exposures. However, cluster 2 was

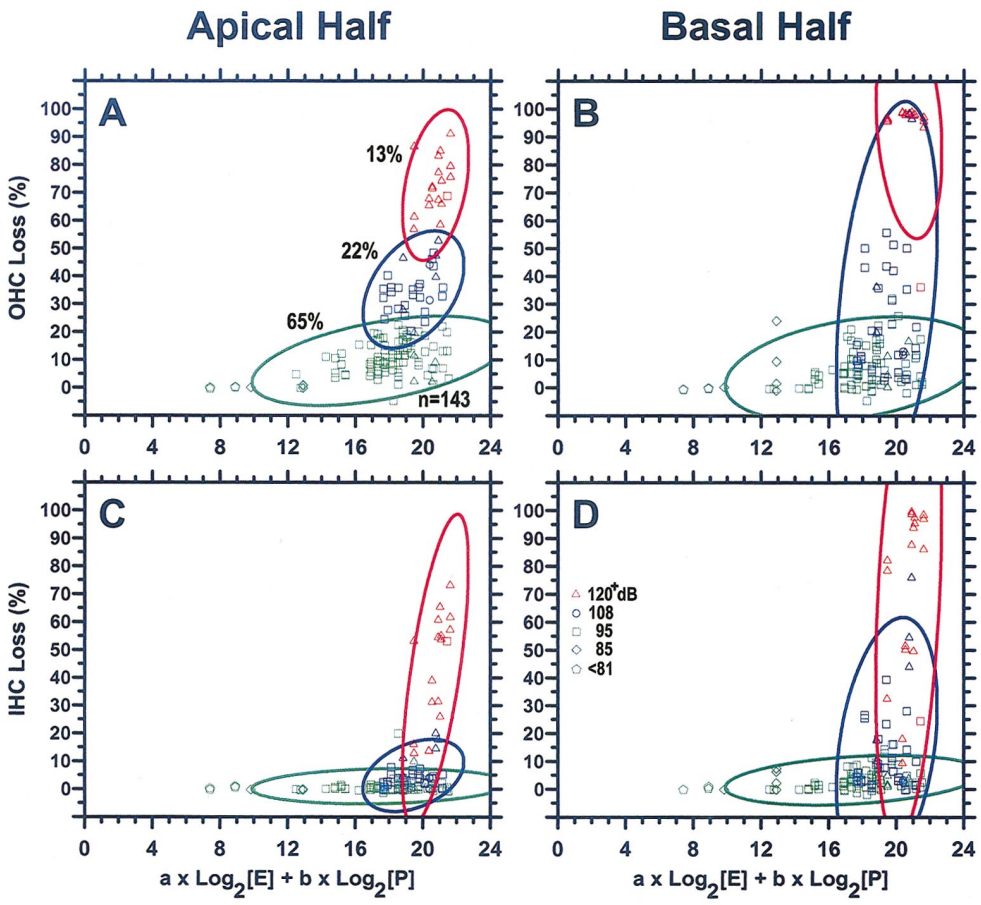


FIG. 7. 0.5 kHz OBN, Chronic: (a) OHC—apex: Poor relation with post-exposure recovery time; cluster 2 (blue) appears with exposures at 95 dB SPL and above; some increase in OHC loss with 120 and 128 dB SPL exposures (triangles). (b) OHC—base: Pattern much like Acute OHC loss in base from 4 kHz OBN. (c) IHC—Apex: Some increase in IHC loss (green) below 120 dB SPL; Much increased IHC loss at 120 and 128 dB SPL. (d) IHC—base: Much like IHC loss in apex. Data from 103 chinchillas.

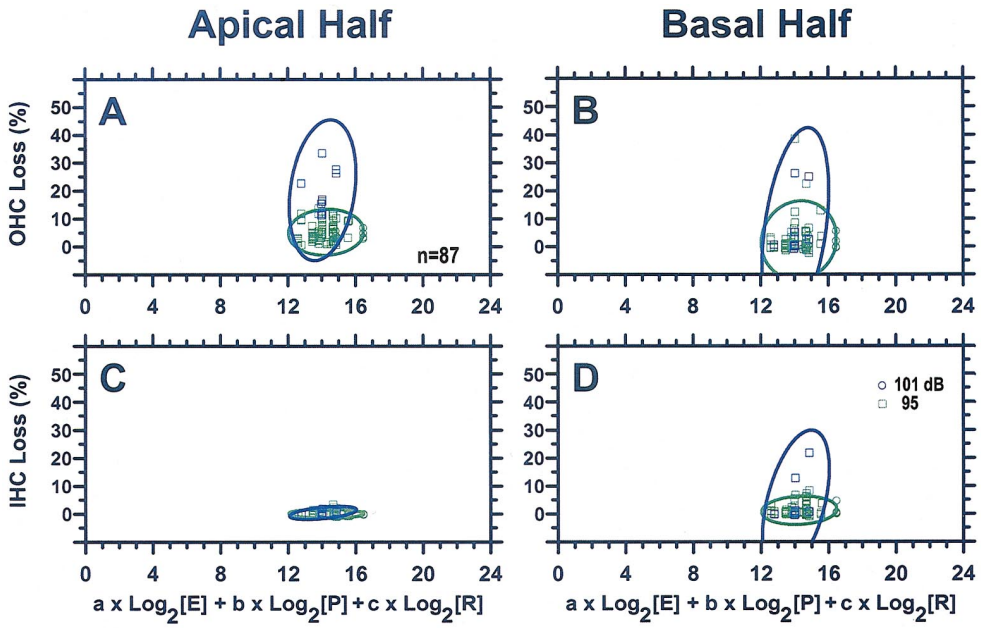


FIG. 8. 0.5 kHz OBN, Interrupted: (a) and (b) OHC—apex and base: Minimal OHC loss, but cluster 2 (blue) evident; slightly less OHC loss with longer rest periods. (c) IHC—apex: Minimal IHC loss. (d) IHC—base: Pattern much like OHC loss in apex and base. Exposures above 101 dB SPL not tested. Data from 59 chinchillas.

TABLE II. 4 kHz OBN, number of ears, percentages, and mean and standard deviation ( ) for cluster centers.

	No. Ears (%) [animals]	Energy+recovery+ rest	Basal OHC	Basal IHC	Apical OHC	Apical IHC
Acute subset						
Cluster 1	60 (67) [37]	5.56 (1.42)	1.62 (3.25)	0.67 (1.46)	−0.04 (0.92)	0.02 (0.55)
Cluster 2	21 (23) [15]	6.98 (0.26)	27.04 (12.88)	2.15 (1.90)	1.60 (2.52)	−0.04 (0.35)
Cluster 3	9 (10) [7]	7.90 (0.69)	91.92 (4.93)	8.49 (8.23)	18.55 (14.15)	0.54 (1.03)
Total:	90 (100) [59]					
Chronic subset						
Cluster 1	78 (42) [58]	8.73 (1.95)	6.25 (6.88)	2.97 (5.20)	1.48 (3.88)	0.05 (0.61)
Cluster 2	13 (7) [9]	10.52 (0.82)	51.97 (16.39)	12.60 (6.47)	8.79 (9.67)	0.78 (2.78)
Cluster 3	94 (51) [61]	12.31 (1.67)	94.50 (5.11)	48.69 (24.49)	33.12 (16.57)	8.00 (11.65)
Total:	185 (100) [128]					
Interrupted subset						
Cluster 1	46 (100) [30]	14.55 (2.79)	0.92 (2.45)	0.64 (1.71)	−0.36 (1.91)	−0.30 (0.69)
Total:	46 (100) [30]					

evident here, whereas it was not with the 4 kHz OBN. This suggests that critical level for the 0.5 kHz OBN exposure was about 95 dB SPL in chinchillas. In all of the Acute, Chronic, and Interrupted subsets, equivalent hair-cell loss from the 0.5 kHz OBN required more total energy than that for the 4 kHz OBN.

H. Specific doubling rates

Table IV lists the coefficients (doubling rates) determined from the model [Eq. (2)] for hair-cell loss versus total energy, recovery time, and rest for cluster 1. This hair-cell loss was primarily produced by noise exposures that were

TABLE III. 0.5 kHz OBN, number of ears, percentages, and mean and standard deviation ( ) for cluster centers.

	No. Ears (%) [animals]	Energy+recovery+rest	Apical OHC	Apical IHC	Basal OHC	Basal IHC
Acute subset						
Cluster 1	47 (84) [32]	12.42 (2.95)	4.12 (4.63)	0.04 (0.46)	0.75 (2.66)	0.58 (1.23)
Cluster 3	9 (16) [5]	18.93 (0.68)	37.36 (9.33)	2.24 (1.30)	76.87 (38.22)	3.14 (3.38)
Total:	56 (100) [37]					
Chronic subset						
Cluster 1	94 (65) [68]	17.47 (3.03)	9.51 (6.42)	0.86 (2.53)	6.70 (7.48)	3.10 (3.56)
Cluster 2	31 (22) [22]	19.45 (1.14)	35.11 (8.04)	4.76 (4.93)	31.36 (27.19)	15.13 (17.72)
Cluster 3	18 (13) [13]	20.76 (0.71)	72.71 (9.76)	42.91 (20.07)	93.99 (14.54)	69.50 (31.64)
Total:	143 (100) [103]					
Interrupted subset						
Cluster 1	78 (90) [53]	13.56 (1.42)	5.13 (3.29)	0.10 (0.65)	2.07 (5.61)	1.03 (2.02)
Cluster 2	9 (10) [6]	13.54 (1.03)	20.18 (7.67)	0.39 (0.71)	6.51 (10.91)	3.91 (7.85)
Total	87 (100) [59]					

TABLE IV. Model coefficients and Pearson r values ( ) for 4 and 0.5 kHz OBN, cluster 1.

4 kHz OBN	N	Coeff	Basal OHC	Basal IHC	Apical OHC	Apical IHC
Acute	60	$a=$	0.345 (0.366)	0.066 (0.155)	0.025 (0.093)	−0.018 (0.116)
		$k=$	−3.046	−0.218	−0.375	0.266
Chronic	76[a]	$b=$	0.976 (0.532)	0.443 (0.329)	0.284 (0.282)	0.003 (0.023)
Interrupted	46	$c=$	−0.923 (0.772)	−0.276 (0.533)	−0.219 (0.427)	−0.041 (0.313)
0.5 kHz OBN	N	Coeff	Apical OHC	Apical IHC	Basal OHC	Basal IHC
Acute	47	$a=$	0.806 (0.638)	0.007 (0.053)	0.381 (0.526)	−0.007 (0.020)
		$k=$	−8.302	−0.062	−5.124	0.678
Chronic	78[a]	$b=$	0.181 (0.173)	0.154 (0.299)	0.528 (0.424)	0.579 (0.691)
Interrupted	78	$c=$	−0.435 (0.710)	−0.040 (0.347)	−0.127 (0.176)	0.139 (0.485)

[a]Some ears not included due to nonexistent log(0).

below critical level. The left column of regression coefficients was based upon the above-described analysis methods. The coefficients in the other columns to the right were determined in the same way, but the cluster assignments were based upon those for the left column. Thus, the latter were dependent rather than independent coefficients.

For the 4 kHz OBN below critical level, the rates of basal IHC loss and protective effects of rest were usually much smaller than for OHC loss. In general, the rates of OHC and IHC loss were greater during the post-exposure recovery period than during the exposure, particularly in the basal half. The rates for prevention of both OHC and IHC loss with rest were negative as expected and surprisingly large for OHCs in the base. The rates of primary and secondary OHC and IHC loss in the apical OC were negligible. However, the rates of protection from loss with rest were substantial, although not as great as in the basal OC.

For the 0.5 kHz OBN below critical level, the rate of OHC loss in the apical half was the greatest during the exposure and negligible post-exposure. The rate of protection from loss due to rest was much less than for the 4 kHz OBN. The rates of IHC loss in the apical half and protection from loss with rest were negligible. In the basal half, the rates of OHC and IHC loss were similar to that for the basal loss from the 4 kHz OBN. The rates of protection from loss with rest, however, were not nearly as great and the coefficient was positive for IHCs in the base. The latter indicates that, for basal IHCs, the rest schedules used were not protective.

Assuming a linear relation between doubling of total energy and the primary OHC loss shown here, the equal-energy hypothesis would predict that the values of coefficient a from the 4 and 0.5 kHz OBNs would be 10.8 and 2.2, respectively. This is clearly not the case for primary basal OHC loss with the 4 kHz OBN and apical OHC loss with the 0.5 kHz OBN. For both OBNs, OHC loss in the other half of the OC and IHC loss in general, the values of coefficient a were well below 1.0.

I. Other noise-induced pathology

Although not reported in detail here, the data were also encoded for OC wipeouts, focal lesions, and nerve fiber loss. All three of these were more common in the 4 kHz OBN subsets than in the 0.5 kHz OBN subsets. OC wipeouts were rare in the Acute subsets, common in the Chronic subsets, and uncommon in the Interrupted subsets. Focal lesions were common in the high-level Acute subsets, much more common in the Chronic subsets, and uncommon in the Interrupted subsets. Nerve fiber loss was associated with substantial IHC loss.

IV. DISCUSSION

A. Hair-cell loss due to aging

The hair-cell loss from aging was very close to that found in an earlier examination which included some of the same cochleae (Bohne *et al.*, 1990), although the regression analysis used here was slightly different. For OHC loss in the basal half of the OC, there appeared to be a slight increase in slope for animals greater than 10 yr of age. However, there were too few older ears to address this issue with any degree of confidence. The present data revealed a slight, but significantly greater loss of OHCs in the apical half compared to the basal half of the chinchilla OC. The reason for this difference is unknown. Higher OHC loss due to aging in the apex of the guinea pig OC has been reported (e.g., Coleman, 1976; Ingham *et al.*, 1999). There is some evidence that these observations might apply to humans (Felder and Schrott-Fischer, 1995). Because the chinchilla has a low rate of hair-cell loss with age, it can be used for long-term noise-exposure studies without having an interaction between noise-induced hair-cell loss and age-related hair-cell loss.

B. Model of noise-induced hair-cell loss

The model for noise-induced hair-cell loss from below-critical-level exposures [Eq. (2)] was constructed as a convenient way to account for related issues and to analyze the cell-loss data. The coefficients are relative to the scaling of the independent variables. Thus, the specific values can only be interpreted with respect to that scaling. A change of scale would require adjusting their magnitudes accordingly. Because the scales for total energy, recovery time, and rest differ, the magnitudes of the coefficients cannot be interpreted relative to each other. In addition, the clustering was dependent upon OHC loss in the base for cochleae exposed to the 4 kHz OBN and in the apex for cochleae exposed to the 0.5 kHz OBN. The coefficients for OHC loss in the other half of the OC and IHC loss in both halves reflect this dependency.

The values of the coefficients would be different with an independent cluster analysis by cell type and OC region.

C. Susceptibility to noise-induced hair-cell loss

The data from moderate-level exposures (nearly all of cluster 1) suggests that all chinchillas were susceptible to noise-induced hair-cell loss during the exposure, but at a low rate. However, above-critical-level exposures produced a much wider range of hair-cell loss. Most animals were also susceptible to these exposures (cluster 2), but some were hypersensitive (cluster 3), while others were resistant (cluster 1).

The values for constant k from the regressions in the Acute groups, cluster 1 for both OBNs, agree with findings in the literature from Mills (1973) involving a 4 kHz OBN and Carder and Miller (1972) involving a 0.5 kHz OBN. Mills determined by extrapolation that the SPL of a 4 kHz OBN had to be greater than 47 dB before a TTS appeared. Also by extrapolation, Carder and Miller indicated that with a 0.5 kHz OBN, the SPL would have to 65 dB before there was a detectable TTS. It is interesting that the difference between the latter two levels is within 3 dB of the added pressure at the eardrum due to ear-canal resonance around 4 kHz. Bohne (1976) determined that continuous 4 kHz OBN exposures at and below 72 dB SPL for 2 or 9 d produced minimal sensory cell loss. For a 0.5 kHz OBN at 75 dB SPL, there was no significant sensory cell loss at 2 d, but about 10% OHC loss in the apex at 9 d.

D. Patterns of hair-cell loss during 4 kHz OBN vs 0.5 kHz OBN and critical level

From the data presented here, it is difficult to determine the exact magnitude of critical level. With the exception of two animals exposed at 92 dB SPL in the 4 kHz OBN Interrupted subset, animals were not exposed at levels between 86 and 108 dB SPL. Thus, critical level for the 4 kHz OBN exposure can only be estimated from the present data. However, the critical level must account for higher energy delivered to the cochlea due to ear canal resonance. Therefore, it is likely that the critical level for the 4 kHz OBN in the chinchilla is not far above 86 dB SPL, perhaps about 90 dB SPL.

For the 0.5 kHz OBN, only four animals in the Interrupted subset were exposed at a level between 95 and 108 dB SPL. All of these eight ears exposed at 101 dB SPL were in cluster 1. The 120 dB SPL level is clearly above critical level and one of two animals exposed at 108 dB SPL clustered with those at 120 dB SPL, indicating that 108 dB SPL is probably above critical level as well. The appearance of cluster 2 in the 95 dB SPL interrupted exposures suggests that the critical level for the 0.5 kHz OBN is below 95 dB SPL in chinchillas. Ward *et al.* (1981) exposed chinchillas for different durations to a 1.4 kHz OBN at different SPLs. These exposures should have put most of the damage in the middle of the chinchilla OC, overlapping both the apical and basal halves. Ward *et al.* found a substantial increase in the total percentage of OHC loss between exposures at 111 and 120 dB SPL after 30 d of recovery. Based on these results, Clark (1991) suggested that in chinchillas, critical level for continuous OBN exposures is about 115 dB SPL. The data reported here indicate that in chinchillas, 115 dB SPL for a continuous OBN is much too high for our broadened definition of critical level.

Bohne (1976) showed that a continuous, 4 kHz OBN exposure of 80 or 86 dB SPL or a 0.5 kHz OBN exposure of 85 dB SPL or greater for 9 d produced substantial hair-cell loss in focal regions in the basal half of the chinchilla OC. These observations are consistent with the results shown here in that clusters 2 and 3 did not appear, regardless of exposure level, until total energy had reached a certain magnitude. Thus, critical level is dependent upon both exposure level and duration. However, for clusters 2 and 3, the 2^7 Pa^2 s difference in onset of hair-cell loss between the 4 and 0.5 kHz OBNs is not accounted for by a 2^5 Pa^2 correction for 15 dB SPL of ear-canal resonance. Therefore, as observed by Bohne and Clark (1982), the base appears to be more sensitive to above-critical-level exposures than the apex.

A number of studies have been conducted in chinchillas to examine the equal-energy hypothesis with respect to broad-band impact noise (e.g., Henderson *et al.*, 1982; 1991; Roberto *et al.*, 1985; Levine *et al.*, 1998). Roberto *et al.* found exposures at and below a peak level of 119 dB SPL produced hearing loss and hair-cell loss that were consistent with the equal-energy hypothesis. However, above 119 dB SPL, hair-cell loss increased significantly. Levine *et al.* found that a peak level of 125 dB SPL was just below critical level for hearing loss. Henderson *et al.* (1982) found that with several exposures having equal energy, their animals did not develop the same amount of threshold shift as would be predicted by the equal-energy hypothesis. Rather, hearing loss increased with peak level. Using peak levels from 107 to 137 dB SPL, Henderson *et al.* (1991) found that their results did not conform to the equal-energy hypothesis in that both hearing loss and hair-cell loss increased more than would be predicted when peak levels were above 119 dB SPL.

Studies have been conducted in chinchillas to compare the consequences of OBN and broad-band impact noise (e.g., Ahroon *et al.*, 1993; Hamernik *et al.*, 2003). Ahroon *et al.* found that exposures that resulted in less than 10 dB PTS or less than 5% total hair-cell loss tended to produce effects that were consistent with the equal-energy hypothesis. On the other hand, the data from higher peak-level exposures were not consistent with the equal-energy hypothesis. There appeared to be an abrupt transition in hair-cell and hearing loss once a certain exposure level was exceeded. Hamernik *et al.* exposed chinchillas to continuous and interrupted OBN exposures at 85–99 dB SPL, impact noise at peak levels of 115–129 dB SPL, and combinations of the two. They did not address the critical-level issue, but they stated that their results were not consistent with the equal-energy hypothesis.

E. Primary versus secondary hair-cell loss and critical level

Bohne and Harding (2000) found that most of the OHC loss in the base occurs post-exposure from a 4 kHz OBN presented above critical level. One cause of hair-cell loss in

the noise-damaged OC is intermixing of cochlear fluids through the damaged reticular lamina (Ahmad *et al.*, 2003). Until the results of the present study, it was not clear that secondary hair-cell loss dominates the loss from 4 kHz OBN exposures below critical level. It was somewhat unexpected that most of the OHC loss from a 0.5 kHz OBN below critical level occurred during the exposure. However, it has been clearly shown that the patterns of hair-cell loss in the apex and base are different (e.g., Bohne and Harding, 2000).

Some researchers speculated that critical level represents a transition from noise-induced hair-cell loss to mechanical damage or acoustic trauma (e.g., Ward, 1981; Levine *et al.*, 1998). The classic definition of ''critical level'' is the intensity that produces irreversible damage in the OC no matter how brief the exposure. This level is generally thought to represent the ''elastic limit'' that, when exceeded, results in immediate rupture of the OC. However, the primary loss in clusters 2 and 3 appeared at exposure levels well below those commonly associated with mechanical damage and the secondary loss in these clusters increased substantially during post-exposure recovery. These results suggest that the definition of critical level should not be limited to the threshold for mechanical damage. The definition of critical level should be expanded to include the level at which substantial secondary hair-cell loss occurs post-exposure.

F. Rest during the exposure and critical level

In the present study, rest during moderate-level exposures was protective for hair-cell loss; more for the 4 kHz OBN than the 0.5 kHz OBN. Bohne *et al.* (1985; 1987) and Clark *et al.* (1987) examined the effect of rest during the exposure on hearing loss and hair-cell loss. They found that interrupted exposures produced less TTS and PTS and less hair-cell loss than continuous exposures with equal energy. Ward (1991) examined interrupted exposures with equal energy. He found that rest during the exposure reduced OHC loss more than would be predicted by the equal-energy hypothesis. In a large study using narrow-band impact noise, Hamernik and Ahroon (1998) compared continuous and interrupted exposures with equal energy. They found that rest produced a slightly lower PTS but no difference in sensory-cell loss. Dolan *et al.* (1976) reported that with very high level exposures, rest was not protective. The equal-energy hypothesis predicts that continuous and interrupted exposures with equal total energy would produce the same amount of hair-cell loss. Thus, the results reported here do not support the equal-energy hypothesis.

G. Other species

Examinations of the equal-energy hypothesis have also been conducted in species other than chinchillas. Eldredge and Covell (1958), Salt *et al.* (1981), Goulios and Robertson (1983), and Fredelius *et al.* (1987) looked at this issue in guinea pigs. Eldredge and Covell introduced the equal-energy hypothesis from their results. However, none of the latter three studies produced results that were consistent with the assumptions of the equal-energy hypothesis. The critical level appeared to be about 20 dB SPL higher in the guinea pig compared to the chinchilla (e.g., Eldredge and Covell, 1958). Borg and Engström (1989) looked at the equal-energy hypothesis in rabbits using a 4 kHz OBN at either 85 or 115 dB SPL. They found results similar to those reported for chinchillas. Bies and Hansen (1990) conducted a retrospective study of a large human population. They found that the equal-energy hypothesis was not adequately supported. They suggested that NIHL scales better on the integral of pressure rather than pressure squared. However, this conclusion may have been the result of fitting data from a mixture of both above- and below-critical-level exposures.

H. Critical level in humans

Chinchillas are about 10 dB SPL more sensitive to noise than humans (Eldredge *et al.*, 1973). Thus, critical level in humans is expected to be about 10 dB SPL higher than predicted for chinchillas. Also, humans tend to be exposed to noise on an interrupted schedule. The frequency spectrum of the noises to which humans are exposed varies among a number of sources. These spectrums effect the base and apex differently with low-frequency noise causing hair-cell loss in both the apex and base (e.g., Bohne and Harding, 2000). The observations reported here along with those in the literature indicate that critical level is different depending upon the characteristics of the noise exposure and the location in the OC that the noise has its greatest effect. In addition, critical level differs with continuous and interrupted exposures. Also, critical level appears to be species specific.

V. CONCLUDING REMARKS

The principal results from this study are as follows.

(a) Primary hair-cell loss is related to total energy of the exposure, provided that the intensity of the exposure is below critical level. However, the specific doubling rates are not consistent with the equal-energy hypothesis. In addition, the magnitude of the relation is noise frequency-spectrum dependent.

(b) The definition of critical level should include the noise intensity at which a substantial amount of secondary hair-cell loss occurs post-exposure. Critical level is species specific and depends upon the frequency spectrum of the noise and the ear-canal resonance of the animal and its susceptibility to noise.

(c) With exposures below critical level, periods of rest during the exposure substantially reduce OHC and IHC loss for the 4 kHz OBN, less so for the 0.5 kHz OBN.

(d) With a 4 kHz OBN, much OHC loss occurs post-exposure. With a 0.5 kHz OBN below critical level, much OHC loss occurs during the exposure while much of the loss from above-critical-level exposures occurs post-exposure. For both OBNs, most IHC loss occurs post-exposure.

(e) Hair-cell loss in the base resulting from a 0.5 kHz OBN is strikingly similar to that due to a 4 kHz OBN, but requires more energy, in part due to the lack of ear-canal resonance at 0.5 kHz.

Nordmann *et al.* (2000) found that cochleae with TTS have a completely different histopathological appearance than those with PTS and concluded that TTS is partially protective against PTS. With a damaging noise exposure, much hair-cell loss occurs post-exposure. Thus, there is a window of opportunity to administer treatment to prevent or minimize noise-induced secondary hair-cell loss, and therefore, NIHL. However, such treatment should be directed at preventing hair-cell loss and PTS, not TTS.

ACKNOWLEDGMENTS

This work was supported by NINCDS, NIDCD, NIOSH, NOHR, DRF, the Department of Otolaryngology, and the Central Institute for the Deaf. The manuscript contents are solely the responsibility of the authors and do not necessarily represent the official views of these granting agencies. Contributions to data collection were made by numerous colleagues, residents, students, and technicians. The authors are grateful to Dr. William W. Clark, Dr. James D. Miller, and Dr. Alec N. Salt for their critical reviews of the manuscript and many thoughtful and helpful comments.

Ahmad, M., Bohne, B. A., and Harding, G. W. (**2003**). ''An in vivo tracer study of noise-induced damage to the reticular lamina,'' Hear. Res. **175**, 82–100.

Ahroon, W. A., Hamernik, R. P., and Davis, R. J. (**1993**). ''Complex noise exposures: An energy analysis,'' J. Acoust. Soc. Am. **93**, 997–1006.

American National Standards Institute (**1996**). ''Determination of occupational noise exposure and estimation of noise-induced hearing impairment,'' ANSI **S3.44**, 28 pp.

Bies, D. A., and Hansen, C. H. (**1990**). ''An alternative mathematical description of the relationship between noise exposure and hearing loss,'' J. Acoust. Soc. Am. **88**, 2743–2754.

Bohne, B. A. (**1976**). ''Safe level for noise exposure?,'' Ann. Otol. Rhinol. Laryngol. **85**, 711–724.

Bohne, B. A., Bozzay, D. G., and Harding, G. W. (**1986**). ''Interaural correlation in normal and traumatized cochleas: Length and sensory cell loss,'' J. Acoust. Soc. Am. **80**, 1729–1736.

Bohne, B. A., and Clark, W. W. (**1982**). ''Growth of hearing loss and cochlear lesion with increasing duration of noise exposure,'' in *New Perspectives on Noise-induced Hearing Loss*, edited by R. P. Hamernik, D. Henderson, and R. Salvi (Raven, New York), pp. 283–302.

Bohne, B. A., Gruner, M. M., and Harding, G. W. (**1990**). ''Morphological correlates of aging in the chinchilla cochlea,'' Hear. Res. **48**, 79–92.

Bohne, B. A., and Harding, G. W. (**2000**). ''Degeneration in the cochlea after noise damage: Primary versus secondary events,'' Am. J. Otolaryngol. **21**, 505–509.

Bohne, B. A., Yohman, L., and Gruner, M. M. (**1987**). ''Cochlear damage following interrupted exposure to high-frequency noise,'' Hear. Res. **29**, 251–264.

Bohne, B. A., Zahn, S. J., and Bozzay, D. G. (**1985**). ''Damage to the cochlea following interrupted exposure to low frequency noise,'' Ann. Otol. Rhinol. Laryngol. **94**, 122–128.

Borg, E., and Engström, B. (**1989**). ''Noise level, inner hair cell damage, audiometric features, and equal-energy hypothesis,'' J. Acoust. Soc. Am. **86**, 1776–1782.

Carder, H. M., and Miller, J. D. (**1972**). ''Temporary threshold shifts from prolonged exposure to noise,'' J. Speech Hear. Res. **15**, 603–623.

Clark, W. W. (**1991**). ''Recent studies of temporary threshold shift (TTS) and permanent threshold shift (PTS) in animals,'' J. Acoust. Soc. Am. **90**, 155–163.

Clark, W. W., and Bohne, B. A. (**1978**). ''Animal model for the 4-kHz tonal dip,'' Ann. Otol. Rhinol. Laryngol. Suppl. **51**, 1–16.

Clark, W. W., Bohne, B. A., and Boettcher, F. A. (**1987**). ''Effect of periodic rest on hearing loss and cochlear damage following exposure to noise,'' J. Acoust. Soc. Am. **82**, 1253–1264.

Coleman, J. W. (**1976**). ''Hair cell loss as a function of age in the normal cochlea of the guinea pig,'' Acta Oto-Laryngol. **82**, 33–40.

Dolan, T. R., Murphy, R. J., and Ades, H. W. (**1976**). ''A comparison of the permanent deleterious effects of intense noise on the chinchilla resulting from either continuous or intermittent noise,'' in *Effects of Noise on Hearing*, edited by D. Henderson, R. P. Hamernik, D. S. Dosanjh, and J. H. Mills (Raven, New York), pp. 327–340.

Eldredge, D. H., and Covell, W. P. (**1958**). ''A laboratory method for the study of acoustic trauma,'' Laryngoscope **68**, 465–477.

Eldredge, D. H., Miller, J. D., Mills, J. H., and Bohne, B. A. (**1973**). ''Behavioral, physiological and anatomical studies of threshold shifts in animals,'' Proceedings of the International Congress on Noise as Public Health Problem, US EPA, publication No. 550/9-73-008, pp. 237–255.

Felder, E., and Schrott-Fischer, A. (**1995**). ''Quantitative evaluation of myelinated nerve fibers and hair cells in cochleae of humans with age-related high-tone hearing loss,'' Hear. Res. **91**, 19–32.

Fredelius, L., Johansson, B., Bagger-Sjoback, D., and Wersall, J. (**1987**). ''Qualitative and quantitative changes in the guinea pig organ of Corti after pure tone acoustic overstimulation,'' Hear. Res. **30**, 157–167.

Goulios, H., and Robertson, D. (**1983**). ''Noise-induced cochlear damage assessed using electrophysiological and morphological criteria: An examination of the equal energy principle,'' Hear. Res. **11**, 327–341.

Hamernik, R. P., and Ahroon, W. A. (**1998**). ''Interrupted noise exposures: Threshold shift dynamics and permanent effects,'' J. Acoust. Soc. Am. **103**, 3478–3488.

Hamernik, R. P., Qiu, W., and Davis, B. (**2003**). ''Cochlear toughening, protection and potentiation of noise-induced trauma by non-Gaussian noise,'' J. Acoust. Soc. Am. **113**, 969–976.

Harding, G. W., Bohne, B. A., and Ahmad, M. (**2002**). ''DPOAE level shifts and ABR threshold shifts compared to detailed analysis of histopathological damage from noise,'' Hear. Res. **174**, 158–171.

Hartigan, J. A., and Wong, M. A. (**1979**). ''A K-means clustering algorithm: Algorithm AS 136,'' Appl. Stat. **31**, 625–642.

Henderson, D., Salvi, R. J., and Hamernik, R. P. (**1982**). ''Is the equal energy rule applicable to impact noise?,'' Scand. Audiol. Suppl. **16**, 71–82.

Henderson, D., Subramaniam, M., Gratton, M. A., and Saunders, S. S. (**1991**). ''Impact noise: The importance of level, duration, and repetition rate,'' J. Acoust. Soc. Am. **89**, 1350–1357.

Ingham, N. J., Comis, S. D., and Withington, D. J. (**1999**). ''Hair cell loss in the aged guinea pig cochlea,'' Acta Oto-Laryngol. **119**, 42–47.

Levine, S., Hofstetter, P., Zheng, X. Y., and Henderson, D. (**1998**). ''Duration and peak level as co-factors in hearing loss from exposure to impact noise,'' Scand. Audiol. Suppl. **48**, 27–36.

Mills, J. H. (**1973**). ''Temporary and permanent threshold shifts produced by nine-day exposures to noise,'' J. Speech Hear. Res. **16**, 426–438.

Nordmann, A. S., Bohne, B. A., and Harding, G. W. (**2000**). ''Histopathological differences between temporary and permanent threshold shift,'' Hear. Res. **139**, 13–30.

Roberto, M., Hamernik, R. P., Salvi, R. J., Henderson, D., and Milone, R. (**1985**). ''Impact noise and the equal-energy hypothesis,'' J. Acoust. Soc. Am. **77**, 1514–1520.

Salt, A. N., Konishi, T., Cook, R. O., and Akay, A. (**1981**). ''Comparison between the effects of continuous and impact noise on cochlear potentials in guinea pigs,'' J. Acoust. Soc. Am. **69**, 1746–1752.

von Bismarck, G. (**1967**). ''The sound pressure transformation function from free-field to the eardrum of chinchilla,'' MS thesis, Massachusetts Institute of Technology.

Ward, W. D. (**1973**). ''Susceptibility to TTS and PTS,'' Proceedings of the International Congress on Noise as Public Health Problem, US EPA, publication No. 550/9-73-008, pp. 281–292.

Ward, W. D. (**1991**). ''The role of intermittence in TTS,'' J. Acoust. Soc. Am. **90**, 164–169.

Ward, W. D., Santi, P. A., Duvall, III, A. J., and Turner, C. W. (**1981**). ''Total energy and critical intensity concepts in noise damage,'' Ann. Otol. Rhinol. Laryngol. **90**, 584–590.

Peripheral compression as a function of stimulus level and frequency region in normal-hearing listeners

David A. Nelson[a] and Anna C. Schroder
Clinical Psychoacoustics Laboratory, Department of Otolaryngology, University of Minnesota, Minneapolis, Minnesota 55455

(Received 6 April 2003; revised 22 January 2004; accepted 26 January 2004)

Fixed-probe-level temporal masking curves (TMCs) were obtained from normal-hearing listeners at probe frequencies between 250 and 8000 Hz. The short probe tones were fixed in level (~10-dB SPL). The level of the preceding forward masker was adjusted to obtain masked threshold as a function of the time delay between masker and probe. These *isoresponse* TMCs were obtained for an on-frequency masker, where the masker frequency (Fm) and probe frequency (Fp) were the same, and for an off-frequency masker below the probe frequency (Fm = 0.6Fp). Slopes of off-frequency TMCs for probe tones at 250–1000 Hz were steeper than those for probe tones between 2000 and 4000 Hz, supporting the notion that response growth for Fm = 0.6Fp at lower probe frequencies is not linear. Therefore, a group average off-frequency TMC slope, for probe frequencies between 2 and 4 kHz, was used to calculate response growth at every probe frequency. Input/output response growth curves were derived from the TMCs, and response growth rates were calculated as a function of the masker level in individual ears. At any particular probe frequency, response growth rates varied with input level, from near 1.0 at low input levels, to <0.2 at mid levels, and back to near 1.0 at levels above 80-dB SPL. It was concluded that compression is equally strong at low and high frequencies as it is at mid frequencies.
[DOI: 10.1121/1.1689341]

PACS numbers: 43.66.Ba, 43.66.Dc, 43.66.Mk [NFV] Pages: 2221–2233

I. INTRODUCTION

The normal auditory system exhibits strong nonlinearities that are associated with excellent sensitivity and sharp tuning. These nonlinearities are thought to originate in the outer hair cells of the cochlear partition. Physiological measures indicate that strong nonlinearities exist in basilar membrane motion (Rhode and Robles, 1974; Robles *et al.*, 1986; Murugasu and Russell, 1995; Ruggero *et al.*, 1997; Russell and Nilsen, 1997; Rhode and Recio, 2000) as well as in the growth of the neural response (Yates, 1990; Yates *et al.*, 1990; Cooper and Yates, 1994). These nonlinearities appear to vary with position along the cochlear partition, with stronger nonlinearities in the base of the cochlea, which is excited most easily by high frequencies, than in the apex of the cochlea, which is sensitive to low frequencies. Although very few direct measurements of basilar membrane motion have been made in the apex (Cooper and Rhode, 1995; Rhode and Cooper, 1996; Khanna and Hao, 1999; Zinn *et al.*, 2000), inferences from neural responses suggest that the nonlinearities in the apex are less strong (Cooper and Yates, 1994). Psychophysical evidence also suggests that cochlear nonlinearities are dependent upon the frequency region at which they are measured (for review, see Hicks and Bacon, 1999a). Some of that evidence includes attempts to specify response growth rates from slopes of nonsimultaneous masking growth functions (Moore *et al.*, 1999; Hicks and Bacon, 1999a; Hicks and Bacon, 1999b; Plack and Oxenham, 2000). The results suggest that response growth rates are dependent upon the frequency region in the cochlea, with steeper growth rates reflecting nonlinearities that are less strong at lower characteristic frequencies.

Nelson *et al.* (2001) demonstrated that the slope-ratio procedure, based on forward-masking growth functions obtained in quiet, may be influenced by the spread of excitation toward higher frequencies. The influence can be greater for an off-frequency masker than for an on-frequency masker. Furthermore, the effects of that spread of excitation may vary with frequency region. Consequently, Nelson *et al.* developed a new procedure for estimating response growth from isoresponse Temporal Masking Curves (TMCs),[1] which they demonstrated, are not significantly affected by the spread of excitation toward higher-frequency regions. TMCs describe the masker levels required to just forward mask a fixed-level probe tone as a function of the time delay between the masker and the probe (Nelson and Freyman, 1987). The TMC requires a constant response at the output of an auditory filter, which produces a constant amount of forward-masked threshold shift (masking) at the probe frequency. The input level to that filter is adjusted to maintain the required response as a function of the time delay between the masker and probe. A fixed-level probe tone is chosen at some frequency and the masker level, at some masker frequency, is adjusted until sufficient forward masking is produced to just mask the probe. As time delay is increased, the amount of forward masking decreases, therefore the masker level (input to the auditory filter) must be increased to maintain the same amount of forward masking. The procedure derives input/

[a]Send correspondence to David A. Nelson, Ph.D., U/M Clinical Psychoacoustics Lab., 38286 Kost Trail, North Branch, Minnesota 55056-6793. Phone: (651) 674-4224; fax: (651) 674-4224; electronic mail: dan@umn.edu

output response growth curves from the slopes of TMCs for off-frequency and on-frequency maskers. Response growth rates and compression exponents are then calculated from those input/output curves, as a function of masker level.

The rationale behind their procedure is as follows: The masker level required to forward mask a probe tone with an *on-frequency masker* (Fm=Fp), depends both upon the recovery from forward masking that occurs at the probe-frequency place *and* upon the peripheral compression that exists at the probe-frequency place. By way of contrast, the masker level required to forward mask a probe tone with an *off-frequency masker* (Fm≤0.6Fp) depends *only* upon the recovery from forward masking that occurs at the probe-frequency place. This is because the off-frequency masker, which is nearly an octave below the probe frequency, presumably has a linear response at the probe-frequency place in the cochlea (Yates, 1990; Yates *et al.*, 1990; Ruggero, 1992; Nelson and Schroder, 1997; Oxenham and Plack, 1997; Ruggero *et al.*, 1997; Moore and Oxenham, 1998; Rhode and Recio, 2000). Therefore, the ratio of the recovery slopes for the off-frequency masker and the on-frequency masker, given the same change in time delay, reflects the additional masker level required in the on-frequency case to overcome cochlear compression and produce the same amount of forward masking. Thus, the ratio of masker level changes for off-frequency and on-frequency maskers as a function of time delay, provides an estimate of response growth at the probe frequency, which is governed by cochlear compression. A critical assumption for this method of estimating compression is that the response to an off-frequency masker is linear at the probe frequency place. Recently, that assumption has been questioned. Forward masking results that did not rely on this assumption suggest that compression is relatively independent of frequency region (Lopez-Poveda *et al.*, 2003; Plack and Drga, 2003; Plack and O'Hanlon, 2003).

The purpose of this study was to evaluate the dependence of off-frequency TMC slopes on probe frequency region, to obtain compression estimates at different frequency regions using the Nelson *et al.* (2001) procedure, and to compare those estimates with previous psychophysical estimates of peripheral nonlinearity.

II. METHODS

TMCs were obtained for a masker frequency at (Fm=Fp) and below (Fm=0.6Fp) a probe frequency. Probe frequencies were 250, 500, 750, 1000, 2000, 3000, 4000, 6000, or 8000 Hz. Masker and probe durations were 200 and 20 ms at peak amplitude, respectively. All stimuli had 10 ms raised-cosine rise and decay times. During each test session, delay times (between masker offset and probe offset) were tested in the following order: 42, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, and 140 ms. A minimum 2 ms temporal separation (for the 42 ms delay time) between masker offset and probe onset (each at 10% of peak amplitude) was used in order to minimize physical interactions between masker and probe. For each delay time, the probe was fixed at a level that was approximately 10-dB SL and the masker level was adjusted adaptively to reach the masked threshold for the probe.

Pure tones for the masker and probe stimuli were produced and gated digitally by Tucker Davis Technologies (TDT) D–A converters, routed separately through programmable attenuators, added together in an active mixer, and presented monaurally through a TDH-49 earphone mounted in an MX/AR-1 cushion. Subjects were seated in a double-walled sound-treated booth and conveyed their responses to the computer by pressing buttons on a custom response panel.

A three-interval forced-choice (3IFC) adaptive procedure was used to estimate the masker level needed to just mask the fixed-level probe tones and to measure absolute sensitivity thresholds. During each 3IFC trial, a subject was presented with three observation intervals demarcated by lights. The masker was presented in all three intervals and the probe was presented in only one, randomly selected, interval. The subject indicated which interval contained the probe stimulus by pressing one of three response buttons, after which the correct-answer feedback was given.

Masked thresholds were determined using a transformed up–down adaptive procedure. During the first four level reversals, a relatively large step size of 8 dB and a simple up–down stepping rule were used to move into the target masker-level region. Then a 2-dB step size was used for the next two reversals, again with a simple up–down stepping rule. A 2-up, 1-down, stepping rule, still with the 2 dB step size, was then followed for the final six reversals to estimate the masker level corresponding to 71% correct detection of the probe (Levitt, 1971). Masked threshold was estimated as the mean of the masker levels for the final six reversals. The final data points were based on the average of three or more such thresholds. Standard deviations for the final means were typically under 3 dB. Absolute thresholds for all probes and maskers and for the audiometric frequencies were obtained using a similar 3IFC adaptive procedure that concluded with a 2-down, 1-up stepping rule for the last six level reversals. Final thresholds for the probes and maskers were estimated in the same manner, as were the masked thresholds. The audiograms were based on one threshold estimate.

Two groups of subjects participated in these experiments. Group I consisted of five normal-hearing subjects who were tested at each of six probe frequencies (500, 1000, 2000, 3000, 4000, and 6000 Hz). Their data were used to evaluate the frequency dependence of off-frequency TMC slopes (Fm=0.6Fp), and to obtain a group average off-frequency TMC slope that could be used to calculate compression at all probe frequencies. Group II consisted of nine additional normal-hearing subjects who were tested at one or more probe frequencies that were either 250, 500, 750, 1000, 6000, or 8000 Hz. The average off-frequency TMC slope from Group I was used to calculate compression from the data for each member of Group II. All subjects had absolute thresholds that were 15 dB HL or better (ANSI, 1996) for octave frequencies between 250 and 8000 Hz and at all probe and masker frequencies. All subjects received several hours of practice on forward-masking tasks before data collection commenced.

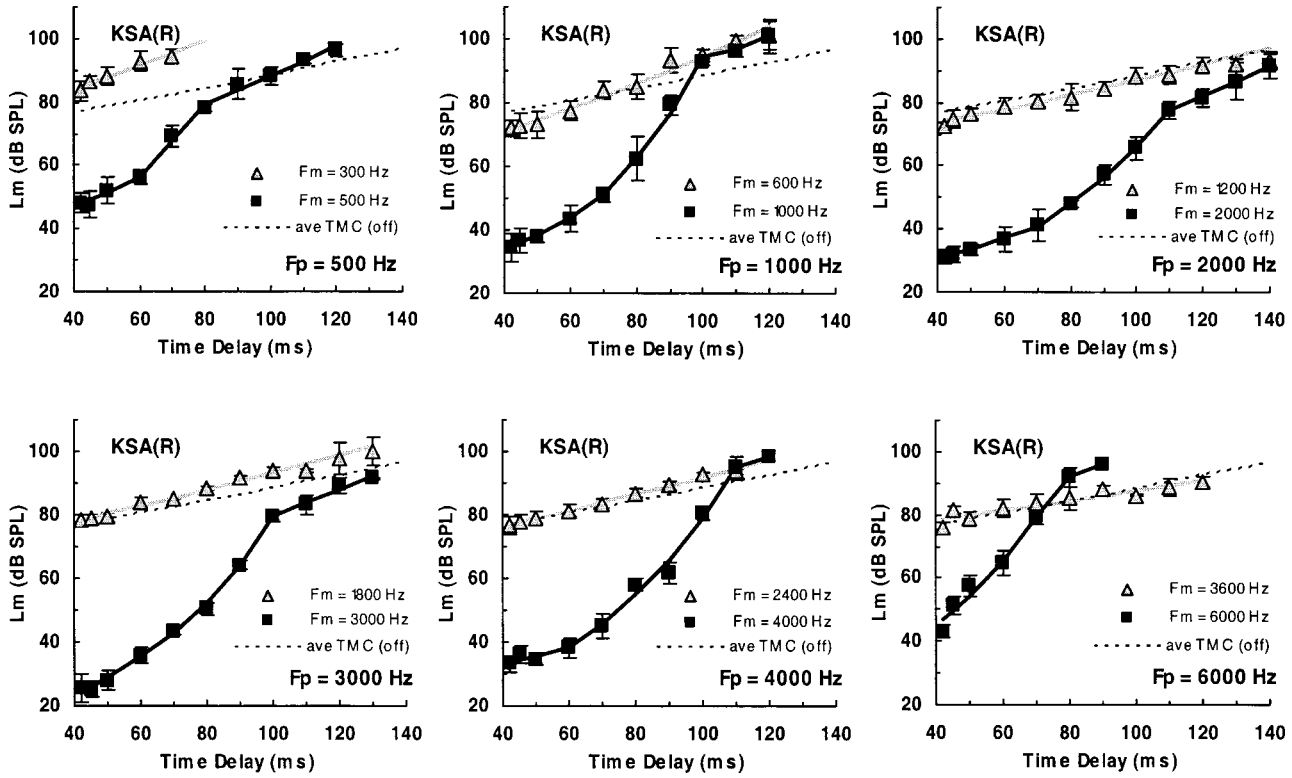


FIG. 1. Temporal masking curves (TMCs) from a normal-hearing subject KSA(R). Each curve shows the masker level required to forward mask a fixed ~10-dB SL probe tone, as a function of the time delay between masker offset and probe offset. Error bars indicate one standard deviation above and below each mean masked threshold. Each panel shows on-frequency (Fm=Fp) and off-frequency (Fm=0.6 Fp) TMCs (black squares and shaded triangles, respectively). The black curves are three-segment exponential fits to the on-frequency data. The shaded curves are single-segment exponential fits to the off-frequency data. The dashed curve shows the average off-frequency TMC for probe frequencies from 2–4 kHz in this subject {ave TMC (off)}. Where the slope of a TMC is steeper than the slope for the ave TMC (off), the assumptions underlying the analysis procedures proposed later, indicate that compression exists.

III. RESULTS

A. Fixed-probe-level (isoresponse) TMCs

TMCs obtained at six probe frequencies from one of the normal-hearing listeners in Group I are shown in Fig. 1. Results for this listener exhibit the general effects seen in all of the subjects. The on-frequency TMCs (black squares) tend to reveal a trisegment shape. Slopes are gradual at low masker levels (short time delays), steeper at mid masker levels (medium time delays), and gradually sloped again at high masker levels (longer time delays). The off-frequency TMCs (gray triangles) maintain a relatively constant slope at all masker levels (all time delays).

Using the procedures introduced by Nelson *et al.* (2001), the relative slopes of on-frequency and off-frequency TMCs can be used to estimate compression as a function of input level. Based on the assumption that the off-frequency masker has a linear response at the probe frequency, the slope of the off-frequency TMC reflects only the recovery from forward masking. As time delay increases, the off-frequency masker must be increased to account for more recovery from forward masking. Usually, with the Nelson *et al.* (2001) procedure, the off-frequency TMC is fit (least squares) with a single exponential to provide a continuous specification of TMC slope, as shown by the shaded curves in Fig. 1. Plack and Drga (2003) suggest that at low probe frequencies (250 Hz), the off-frequency TMC is steeper than at higher frequencies (4000 Hz), and therefore, the off-frequency TMC at low frequencies cannot be used to estimate compression. One explanation for this is that compression extends over a broader *relative* frequency region at low frequencies than at high frequencies. Thus, the off-frequency TMC at low probe frequencies is still influenced by compression. The compression, presumably, increases the TMC slope from what it would be without compression. The fitted off-frequency TMC curves in Fig. 1 (shaded curves) for probe frequencies at 500 and 1000 Hz appear to be steeper than those at higher probe frequencies. To assist this comparison, the dashed line in each panel shows the average of the single exponential fits to the off-frequency TMCs for the 2, 3, and 4 kHz probes obtained from this subject. For this subject, comparisons of individual off-frequency TMCs at each probe frequency (shaded curves) with the average off-frequency TMC (dashed curves) indicates that off-frequency TMCs for probes at 500 and 1000 Hz are slightly steeper than those for probe frequencies at 2000 Hz and higher.

On-frequency TMCs (black squares), in general, exhibit a more complex function. It is assumed that recovery from forward masking is the same with an on-frequency masker as it is with an off-frequency masker. Therefore, if only recovery from forward masking influenced the on-frequency TMC, then the on- and off-frequency TMCs would be parallel but offset to reflect attenuation by the auditory filter.

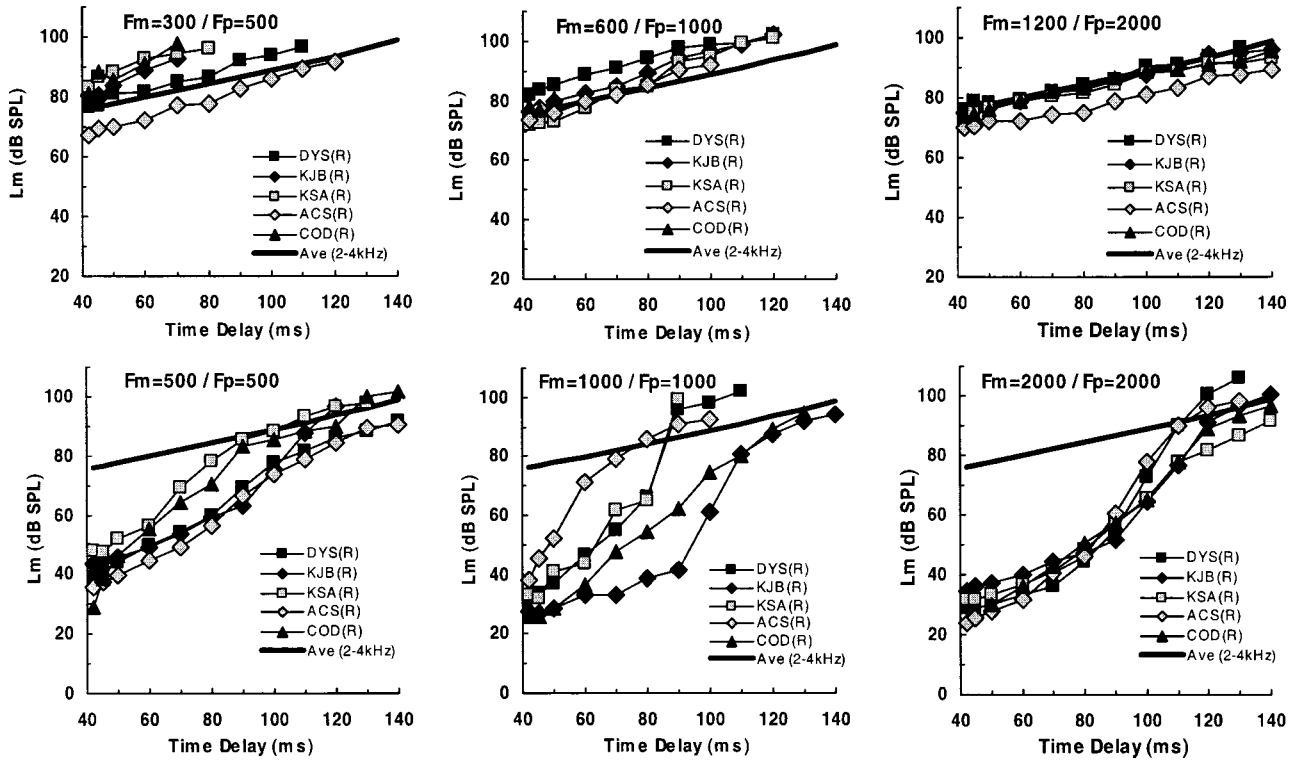


FIG. 2. TMCs from the five Group I subjects for probe frequencies at 0.5, 1.0, and 2.0 kHz. The upper panels show the off-frequency conditions, while the lower panels show the corresponding on-frequency conditions. The *group* average off-frequency TMC {Ave (2–4 kHz)}, shown in each panel by a solid black curve, was based on the average (across the five Group I subjects) of the off-frequency TMCs for probe frequencies from 2000–4000 Hz. This curve was used to derive input/output curves at all probe frequencies.

Thus, the departure of the on-frequency TMC slope from the off-frequency slope reveals the effects of some other process operating on the masker. It is assumed that this process is cochlear compression. As the time delay increases, not only must the masker level increase to account for more recovery from forward masking, but the masker must be increased additionally to overcome the compression imposed on it. The ratio of the slopes of the off-frequency and the on-frequency TMCs at any specific time delay should reveal the magnitude of compression imposed at that particular on-frequency masker level. To aid in the specification of the TMC slope, each of the on-frequency TMCs was fitted with three exponential curves. The fitting procedure was done manually by choosing the number of data points in each segment to minimize differences between fitted and actual data points. However, any fitting procedure that minimizes error would suffice. In general, the fitted curves reveal low-level and high-level segments that are nearly parallel to the off-frequency TMC. The mid-level segment has a steeper slope that reflects the most compression.

Figures 2 and 3 show the off-frequency and on-frequency TMCs obtained from all five subjects in Group I. Each column of panels represents a different probe frequency. On-frequency TMCs are shown in the bottom row of panels in each figure. In general, most on-frequency TMCs exhibit the trisegment form, with individual subject differences in the extent and slope of each segment. Off-frequency maskers are shown in the top row of panels in each figure. All of the off-frequency TMCs were well fit with a single exponential. The solid black curve in each panel shows the *group* average off-frequency TMC, averaged across all five subjects at probe frequencies from 2–4 kHz {Ave (2–4 kHz)}. Notice that the individual off-frequency TMCs from different subjects are well described by the group average off-frequency TMC curve, except at probe frequencies of 500 and 1000 Hz. At these two probe frequencies, some of the individual curves would be better fit with a steeper function.

B. Slopes of off-frequency TMCs by probe frequency

Figure 4 shows the fitted off-frequency TMCs at different probe frequencies, for each of the five subjects in Group I. Those TMCs are compared to the *individual* average off-frequency TMC for probe frequencies from 2000–4000 Hz, in each subject, which is shown by the solid curve {Ave TMC (Off)}. Each panel shows this comparison for a different subject. Notice that the fitted TMCs at 500 and 1000 Hz probe frequencies (unfilled symbols) are steeper than the fitted TMCs at higher probe frequencies (filled symbols). This is the case for four of the subjects. Subject ACS exhibited off-frequency TMCs that had essentially the same slopes across probe frequency. Notice also that, for each subject, the Ave TMC (Off) function describes the off-frequency TMC slopes quite well for probe frequencies above 1000 Hz.

From these comparisons, it is clear that fitted off-frequency TMCs for probe frequencies at 500 and 1000 Hz are steeper than they are at higher probe frequencies, for most subjects. This result suggests that off-frequency maskers for probe frequencies at and below 1000 Hz may not

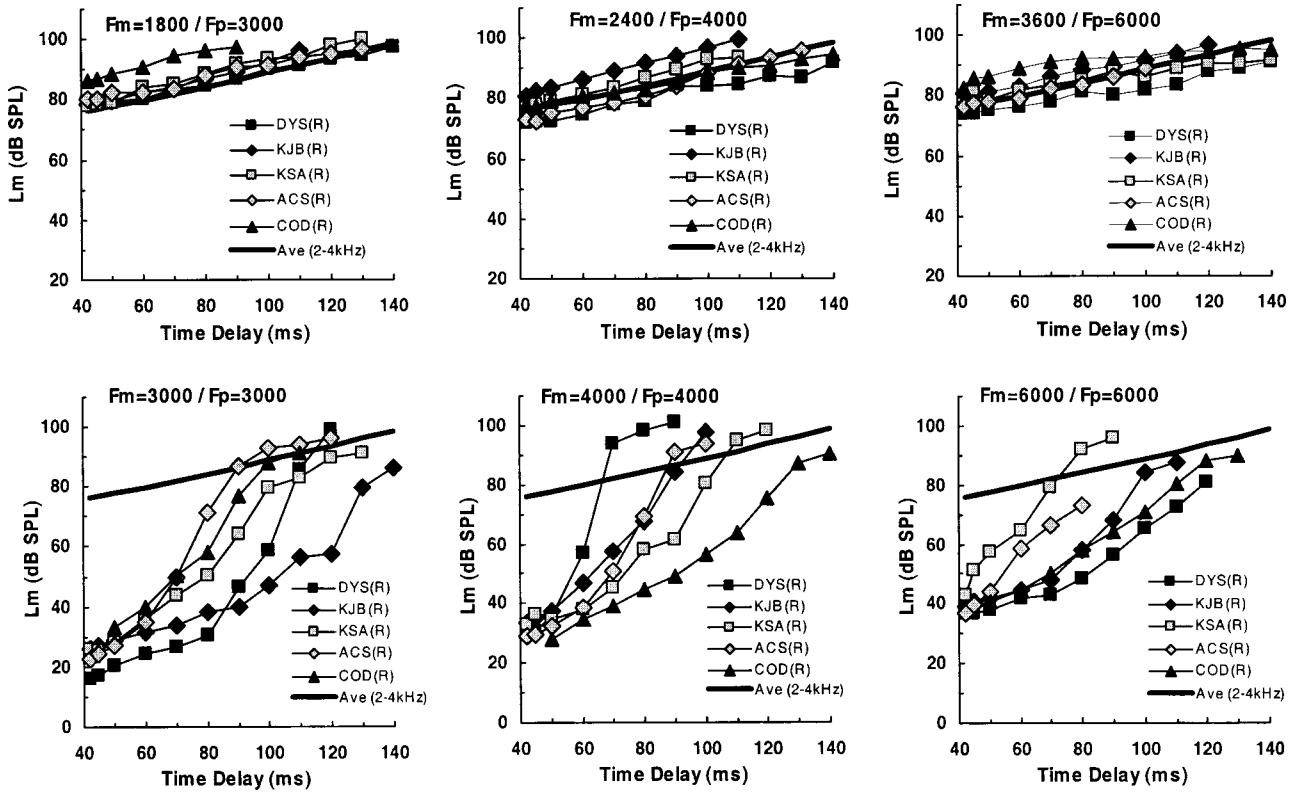


FIG. 3. The same as Fig. 2, but for probe frequencies at 3.0, 4.0, and 6.0 kHz.

exhibit linear response growth at the probe frequency. If this is the case, then the off-frequency TMC slopes for the probe frequencies at 500 and 1000 Hz cannot be used to calculate compression. As suggested by other investigators (Lopez-Poveda *et al.*, 2003; Plack and Drga, 2003), the off-frequency TMC slope from a higher-frequency region, where

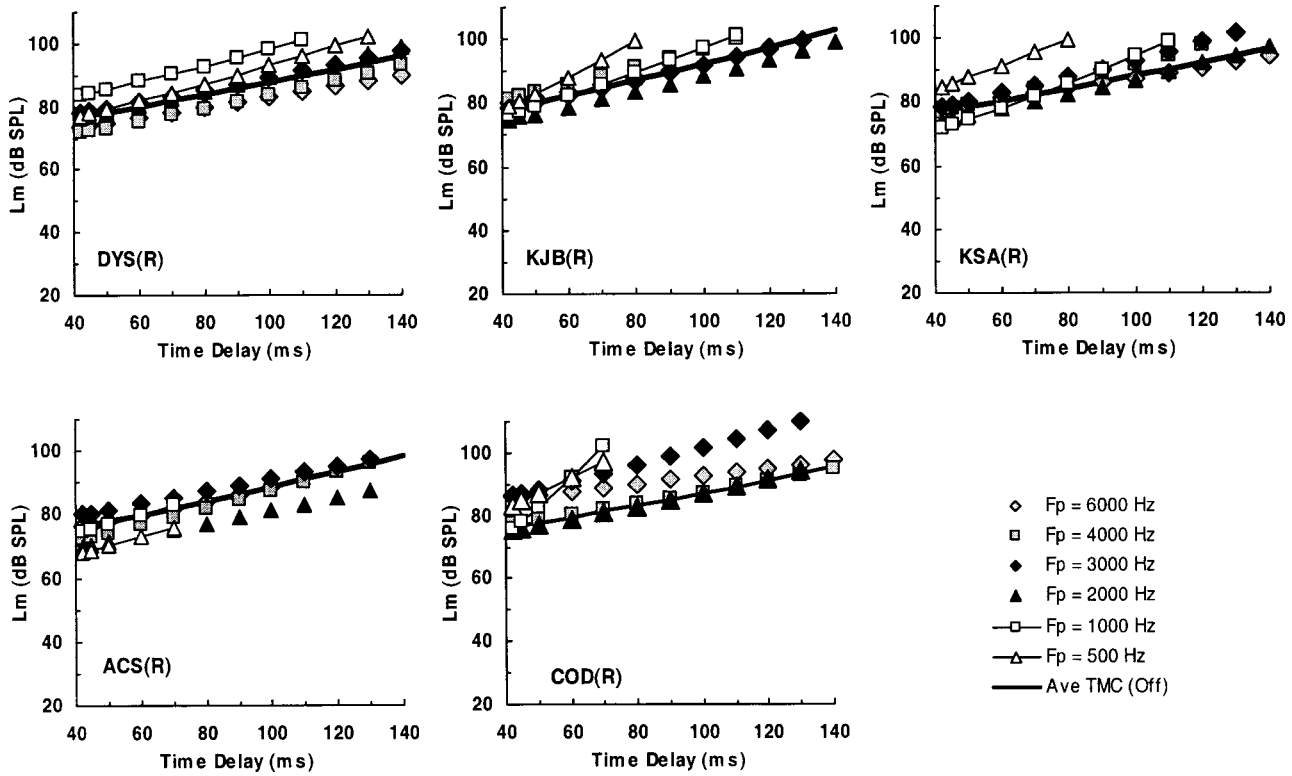


FIG. 4. Off-frequency TMCs (Fm=0.6 Fp) are compared across probe frequencies in the five Group I subjects. The solid black curve in each panel is the average off-frequency TMC at probe frequencies between 2000 and 4000 Hz {ave TMC (Off)} for each individual subject. Probe frequencies at 500 and 1000 Hz, which exhibit steeper off-frequency TMCs, are shown by the open symbols.

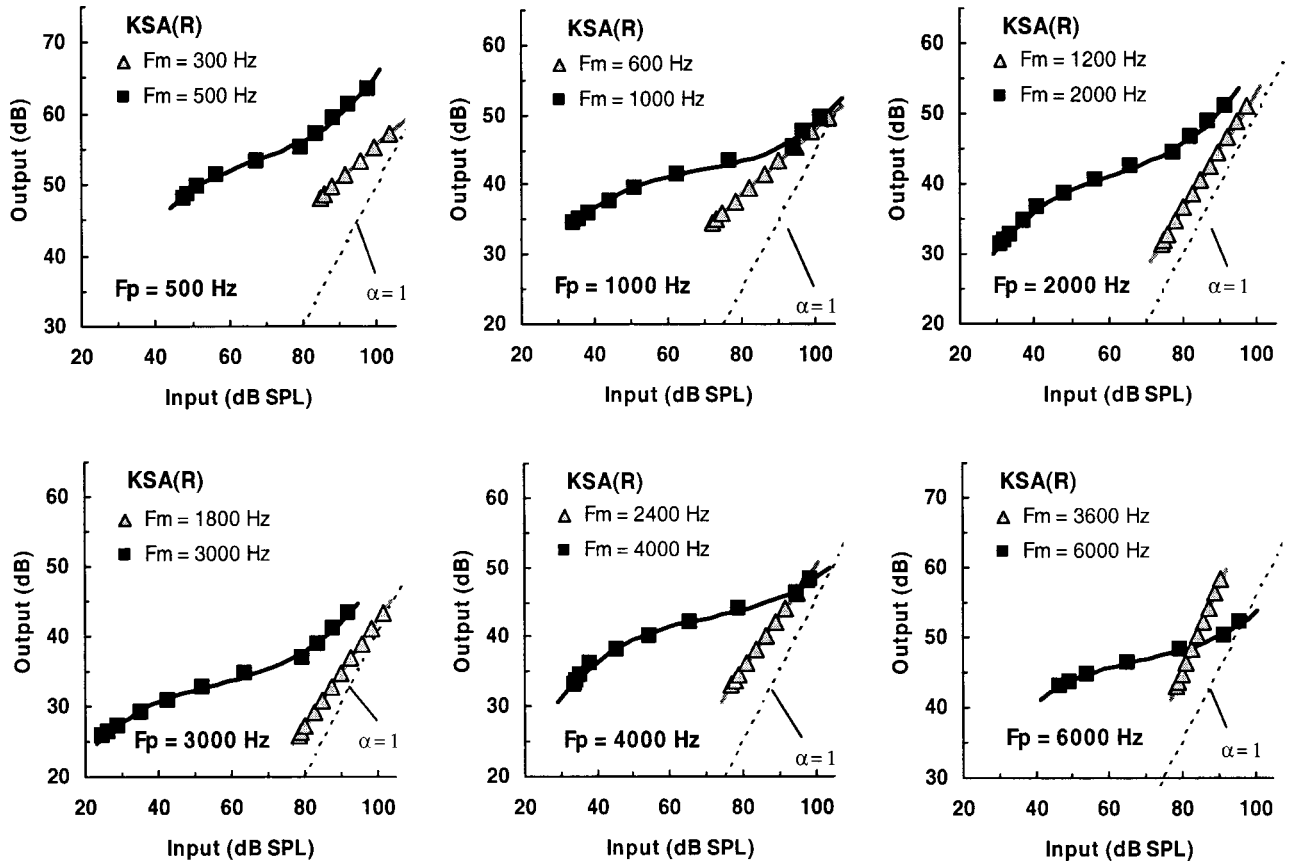


FIG. 5. Input/output curves derived from off-frequency and on-frequency TMCs from subject KSA. On-frequency I/O curves are shown by black squares. Off-frequency I/O curves are shown by shaded triangles. Each panel shows derived I/O curves for a different probe frequency. Note that the origin of the ordinate is 10 dB higher for the 500- and 6000-Hz data. The dotted curve in each graph shows a linear response growth rate ($\alpha=1$), where output (dB) $=\alpha\times$ input (dB) $+K$ (an arbitrary constant to adjust the origin on the graph). The solid curves are third-order polynomial fits to the I/O curves, the first derivatives of which were used to specify the response growth rate as a function of input level.

the linearity assumption is more likely to be valid, can be used instead. For the determination of input/output curves and compression characteristics that follow, we used the *group* average off-frequency TMC (shown by the solid curves in Figs. 2 and 3) as the reference from which to calculate off- versus on-frequency TMC slope ratios that specify compression as a function of the input masker level.

C. Derived I/O curves and compression calculations

Compression was calculated from the TMCs in each subject using the procedures developed by Nelson *et al.* (2001), except for the use of the *group* average 2–4 kHz off-frequency fitted TMC instead of the off-frequency fitted TMC (Fm=0.6Fp) associated with each probe frequency. To illustrate this procedure, Fig. 5 shows input/output (I/O) curves derived from TMCs for one of the Group I subjects (KSA). Input levels are the masker levels corresponding to masked thresholds at different time delays, taken from the fitted on-frequency TMC (e.g., solid black curves in Fig. 1). Corresponding output levels are calculated as the changes in input level with time delay, multiplied by estimates of compression. The estimates of compression are usually specified by the ratios between off-frequency and on-frequency fitted TMC slopes, for consecutive time delays (Nelson *et al.*, 2001). However, given the rationale presented above, the *group* average off-frequency TMC (Fp=2–4 kHz) was used to calculate off- versus on-frequency TMC slope ratios. This also allowed I/O curves to be calculated for the off-frequency conditions at each probe frequency. Thus, in Fig. 5, a linear off-frequency I/O curve with a growth rate of 1.0 ($\alpha=1$) means that the off-frequency TMC for a particular probe frequency had the same slope as the *group* average 2–4 kHz off-frequency TMC curve. Comparisons of the shaded triangles in Fig. 5 with the dotted lines ($\alpha=1$) indicate that off-frequency growth rates at 500 and 1000 Hz were more gradual than at higher probe frequencies.

Our comparisons of off-frequency TMCs at different probe frequencies for Group I subjects revealed consistent differences in off-frequency TMC slopes only at 500 and 1000 Hz. Thus, it is likely that, at those two frequencies, the off-frequency masker response growth was not linear at the probe place, but was subjected to some compression. The effect of this compression is evident in Fig. 4, where the fitted TMCs at 500 and 1000 Hz are steeper than the *individual* average 2–4 kHz off-frequency TMCs in four out of five subjects. This deviation was also apparent in the derived I/O curves for the off-frequency maskers. For example, in Fig. 5, the effects of compression on off-frequency maskers are evident where the derived I/O curves for the off-frequency maskers (shaded triangles) depart from a slope of 1.0 (dashed lines).

The on- frequency I/O curves for subject KSA in Fig. 5

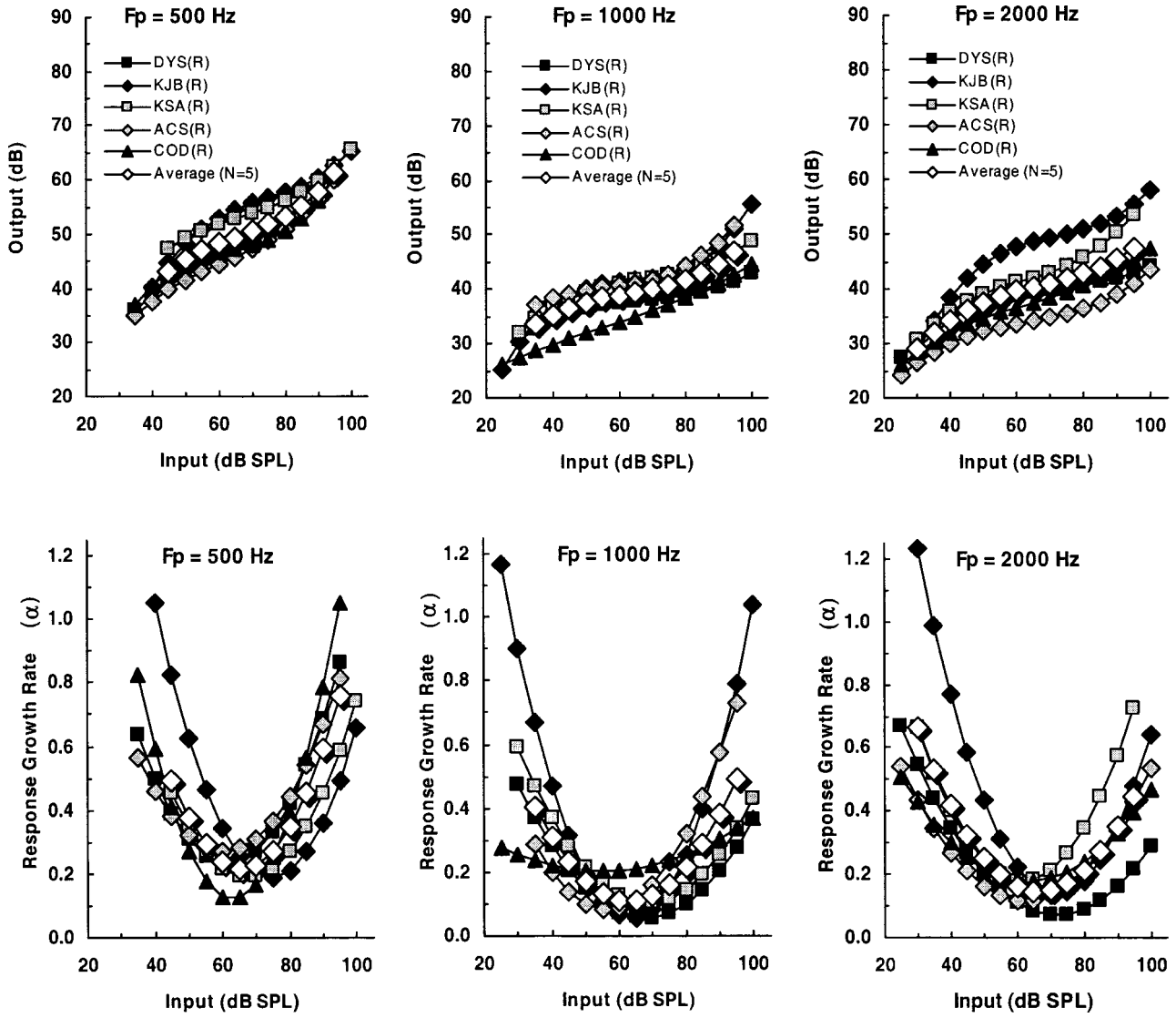


FIG. 6. Input/output response growth curves derived from TMCs (top row), along with response growth rates calculated from the TMCs (bottom row) for the five Group I subjects. Probe frequencies were 500, 1000, and 2000 Hz. Calculations were made at fixed input levels (in 5 dB steps) in order to allow averaging across subjects. Average values, indicated by open symbols, were only calculated at those input levels where data were available from all subjects. The symbols used for the I/O curves in the top row also apply to the response growth rate functions in the bottom row.

(black squares) revealed compression at all six probe frequencies. The general form of every on-frequency I/O curve in this subject was similar, with relatively linear growth for low-level inputs, compressed growth for mid-level inputs, and more linear growth again for high-level inputs. To quantify the magnitude of compression as a function of input level, the derived I/O curves were fitted with a third-order polynomial. In this way the first derivative of the polynomial could be used to specify growth rate (slope of the I/O curve) as a function of input level.

D. Comparisons of I/O curves by frequency region

Similar procedures were applied to the TMCs obtained from all subjects. Then, in order to describe the average result across subjects as a function of input level (dB SPL), the derived I/O curves for the on-frequency conditions were recalculated (in 5 dB steps) from the polynomial fits for each subject using the same exact input levels. This facilitated comparisons among subjects at the same input levels for each probe frequency. It also allowed average I/O curves across subjects to be calculated at each probe frequency to facilitate comparisons of results across frequency regions. Those polynomial fits for the five Group I subjects are shown in Figs. 6 and 7 for different probe frequencies.

The top row of panels in each figure shows the polynomial-fitted I/O curves plotted in 5 dB steps. Averages are shown by unfilled diamonds at each probe frequency. Different probe frequencies are shown in different columns of panels. The bottom row of panels displays the first derivative of the I/O curves, which defines the response growth rate as a function of the input level for each subject. The response growth rate can be interpreted as the compression that exists at any specific input level. All of the growth rate curves are U-shaped, indicating that compression varied continuously with input level. Maximum compression (minimum response growth rate) occurred at moderate input levels (60–70-dB SPL). More linear response growth rates were exhibited at lower and higher input levels.

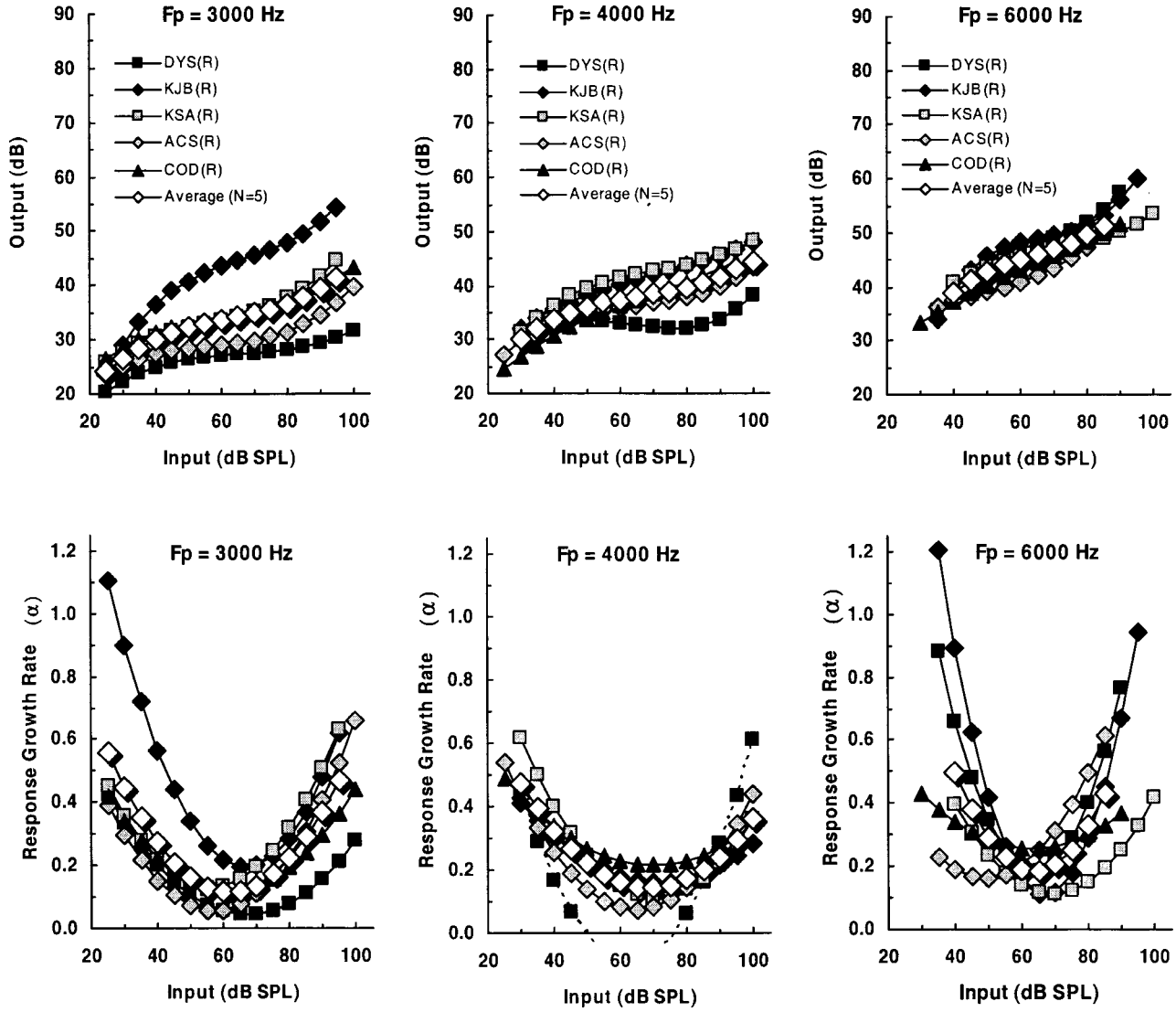


FIG. 7. The same as Fig. 6, but for probe frequencies at 3000, 4000, and 6000 Hz.

Individual differences can be seen among subjects at some frequencies. Those individual differences tended to be larger at lower and higher levels, which can be seen most easily in the U-shaped response growth-rate curves (the bottom row of panels in Figs. 6 and 7). The least variance across subjects occurred in the middle of the input-level range, where compression was the strongest. The negative growth rates at 4000 Hz for subject DYS (black squares with a dashed line) were an artifact of the polynomial fitting procedure; there were no nonmonotonicities in the raw data of the derived I/O curve. The general trends at each probe frequency are summarized by the averaged group curves, shown in Figs. 6 and 7 by the unfilled diamonds. These averages were only calculated at input levels where data existed from all five subjects.

E. I/O Curves from additional frequency regions

The studies by Hicks and Bacon (1999a), Moore *et al.* (1999), and Plack and Oxenham (2000) suggested that the amount of compression in the auditory system varies depending on the frequency region of the cochlea. However, the average minimum response growth rates from the Group I subjects did not appear to be particularly higher at 500 Hz than at higher probe frequencies, as would be expected if there were less compression at lower frequencies. Thus, it became apparent that it would be useful to extend the current investigation of compression to lower and higher probe frequencies than originally planned. Subsequently, data from other laboratories suggested that compression might be independent of probe frequency between 250 and 8000 Hz. Plack and Drga (2003) obtained compression estimates in three subjects at 250 Hz, which were not different from those at 4000 Hz. Lopez-Poveda *et al.* (2003) obtained compression estimates in three subjects at 8000 Hz, which were not different from those obtained at lower frequencies.

Because the Group I subjects were not all available for testing at additional probe frequencies, a slightly different experimental design was implemented. Individual subjects were tested at one or more probe frequencies; no attempt was made to test each subject at all probe frequencies. The final sample included nine subjects, with 5, 7, 7, 6, 5 and 4 subjects tested at probe frequencies of 250, 500, 750, 1000, 6000, and 8000 Hz, respectively. Instead of using off-

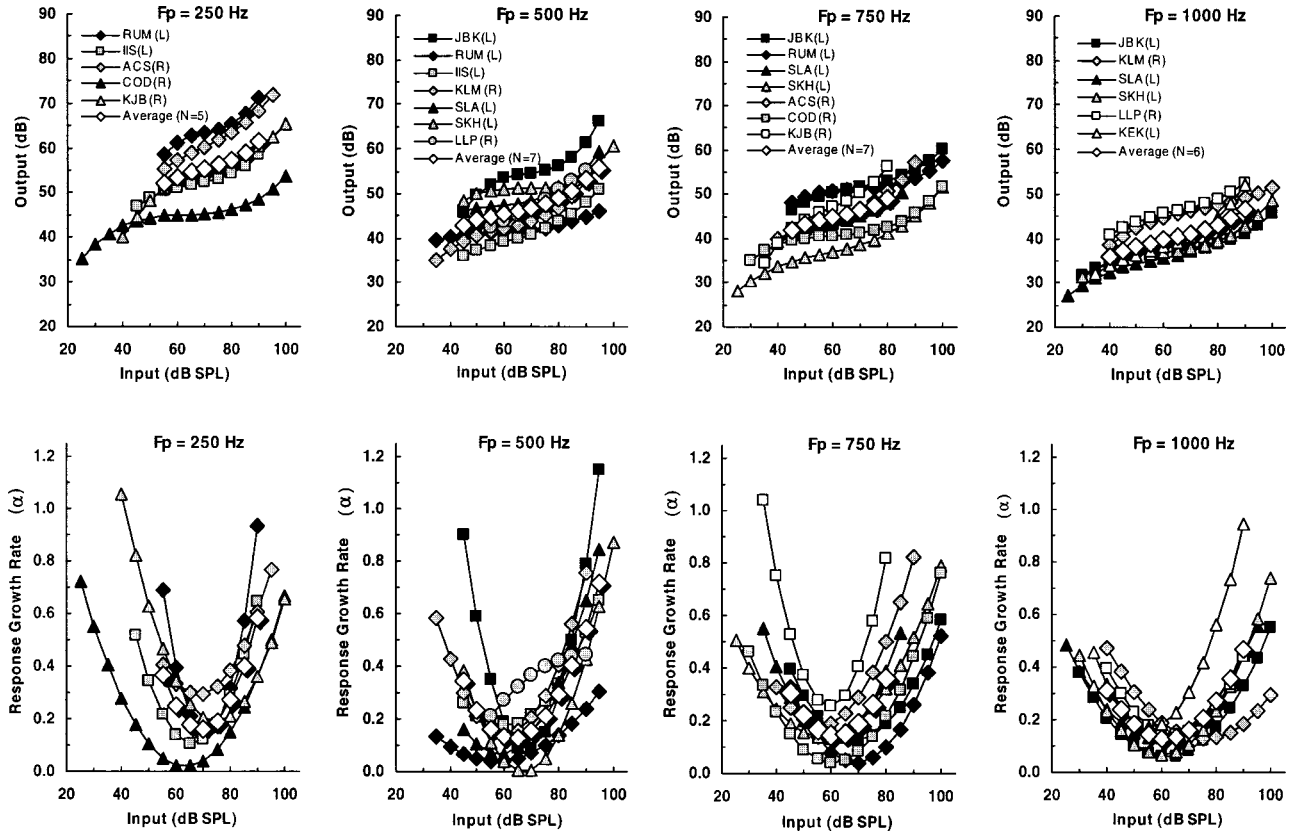


FIG. 8. Input/output response growth curves derived from TMCs (top row), along with response growth rates calculated from the TMCs (bottom row) for Group II subjects. Probe frequencies were 250, 500, 750, or 1000 Hz. Calculations were made at fixed input levels (in 5 dB steps) in order to allow averaging across subjects. Average values, indicated by open diamonds, were only calculated at those input levels where data were available from all subjects. The symbols used for the I/O curves in the top row also apply to the response growth rate functions in the bottom row.

frequency TMCs at each probe frequency, the *group* average 2–4 kHz off-frequency TMC obtained from the Group I subjects was used to calculate compression for these Group II subjects. Although both off-frequency and on-frequency TMCs were obtained from Group II subjects, the results shown here only include derived I/O curves and response growth curves based on the Group II individual on-frequency TMCs and the Group I average 2–4 kHz off-frequency TMC.

The off-frequency TMC slopes from Group II subjects were similar to those obtained from Group I subjects. The 250 Hz off-frequency TMCs were steeper than at higher probe frequencies, but an added complication was encountered with the very high masker levels required to forward mask 250 Hz probe tones. Because probe thresholds were higher at 250 Hz, the masker levels required to mask those probe tones were also higher. Thus, masked thresholds for only the shortest time delays could be obtained before exceeding the limits of the stimulus system at longer time delays. Thus, even if the off-frequency TMC slopes for the 250 Hz probe had been as gradual as at higher probe frequencies, the off-frequency TMC slopes for the 250 Hz probe would have been marginally useful for specifying compression.

Using the same graphical format as in Figs. 6 and 7, Fig. 8 shows derived I/O curves and response-growth estimates obtained from Group II subjects, for probe frequencies between 250 and 1000 Hz. Essentially, the derived I/O curves and response growth-rate curves for Group II behaved the same as for Group I. As with the Group I subjects, all of the growth-rate curves were U-shaped, indicating that compression varied continuously with input level. Maximum compression (minimum response growth rate) occurred at moderate input levels (60–70-dB SPL); response growth rates at lower and higher input levels were more linear. At the 250 Hz probe frequency, the I/O curves exhibited slightly higher output levels, which was due primarily to higher absolute thresholds at 250 Hz than at the higher frequencies. The response-growth rates were just as steep as at higher probe frequencies, but the range over which strong compression occurred was slightly smaller at 250 Hz than at higher frequencies.

Figure 9 shows derived I/O curves and response growth estimates obtained from the Group II subjects who were tested at a probe frequency of 6000 or 8000 Hz. At these high frequencies, I/O curves were similar to those observed at lower probe frequencies, with near linear growth at low and high input levels and compressive growth at mid-input levels. Response growth rate curves were U-shaped with maximum compression (minimum growth rates) at mid-levels. One subject (KSA) was more sensitive at 8000 Hz than the other three subjects, so data from that subject were not included in calculation of the average curve shown in Fig. 9.

To compare results across frequency regions more easily, the across-subject averaged results for Group I and Group II subjects are replotted in Fig. 10. Several general

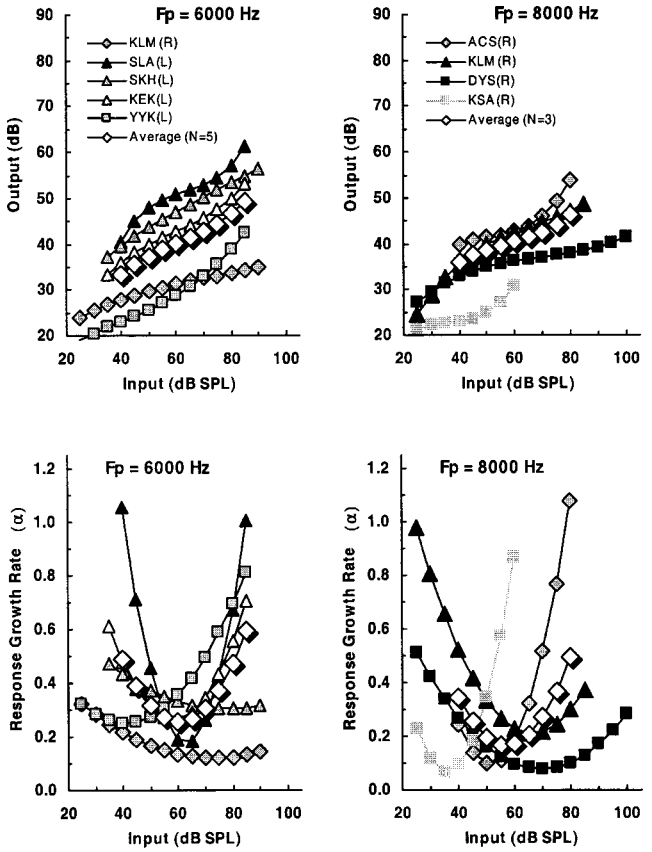


FIG. 9. The same as Fig. 8, but for probe frequencies at 6000 and 8000 Hz. Subject KSA was not included in the average at 8000 Hz.

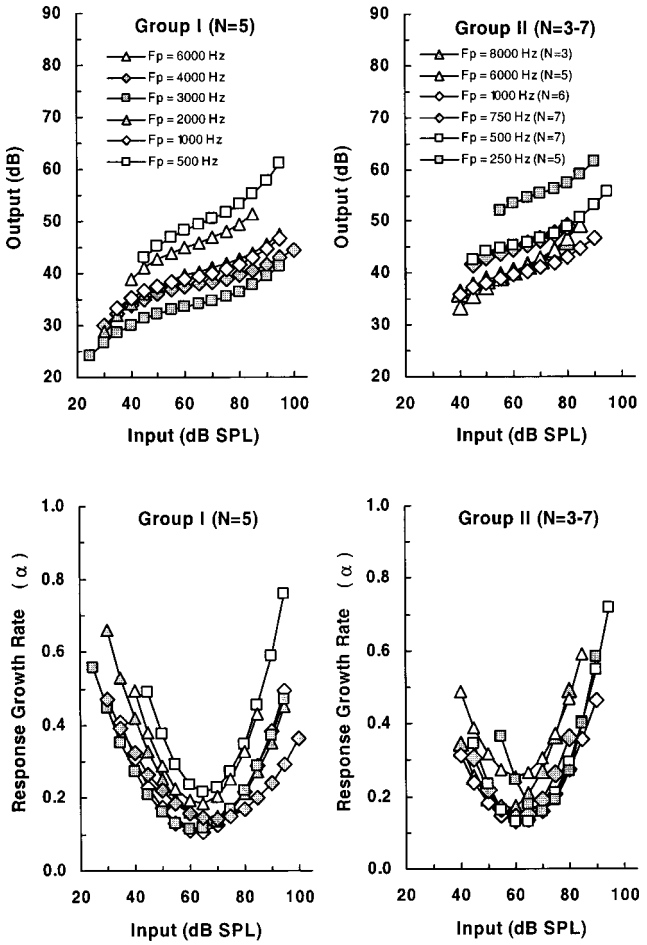


FIG. 10. Across-subject average input/output response growth curves derived from TMCs (top panels), along with average response growth rates (bottom panels) calculated from the TMCs from Group I subjects (left panels) and Group II subjects (right panels). The parameter is probe frequency.

findings are apparent in this figure. Output levels were slightly higher at 250, 500, 750, and 6000 Hz probe frequencies (top panels), primarily because probe thresholds were higher at these frequencies. Since the lowest output level was always the on-frequency masker level at the shortest time delay (42 ms), higher probe thresholds required higher probe levels, which required higher masker levels. Compression was the strongest (lowest response growth rates) for input levels between 60- and 70-dB SPL, regardless of probe frequency. Strong compression was limited to a slightly smaller range of input levels at 250, 500, 6000, and 8000 Hz, which can be seen by their slightly narrower U-shaped response growth-rate functions (bottom panels).

F. Comparisons of maximum compression by frequency region

In order to compare compression across probe frequency regions, the minimum response growth rate, estimated across all input levels, was estimated for each subject at each probe frequency (minimums in the growth-rate versus input-level curves in Figs. 6, 7, 8, and 9). The average minimum response growth-rate (across subjects) then provides an estimate of the maximum compression seen at each probe frequency. Figure 11 shows the average minimum response growth rate for Group I subjects (unfilled squares) and Group II subjects (unfilled triangles) as a function of probe frequency.

An examination of the results in Fig. 11 (unfilled squares and triangles) indicates that the average minimum growth rate does not vary noticeably across probe frequency region. The average minimum growth rate at each probe frequency for Group I subjects (unfilled squares) varied between 0.10 and 0.20 (see Table I). A single-factor ANOVA of Group I results showed that there were no significant differences in average minimum growth rates across probe frequency ($p > 0.10$). The average minimum growth rate at each probe frequency from Group II subjects (unfilled triangles) varied between 0.10 and 0.22 (see Table I). When the two groups were combined, the average response growth rate ranged between 0.11 and 0.20 at different probe frequencies, with a grand mean of 0.14 across subjects and probe frequencies (solid line in Fig. 11). Regression analysis of the averaged minimum response growth rates yielded a nonsignificant trend across probe frequency ($p > 0.69$; $R^2 = 0.02$). Thus, the average minimum growth rate appears to be independent of probe frequency. At its maximum along the I/O curve, compression is equally strong at all frequency regions.

IV. DISCUSSION

The major finding of this study, that compression is relatively constant across probe frequency regions, differs from some previous psychophysical estimates of compression that utilized on- versus off-frequency growth-of-masking slope

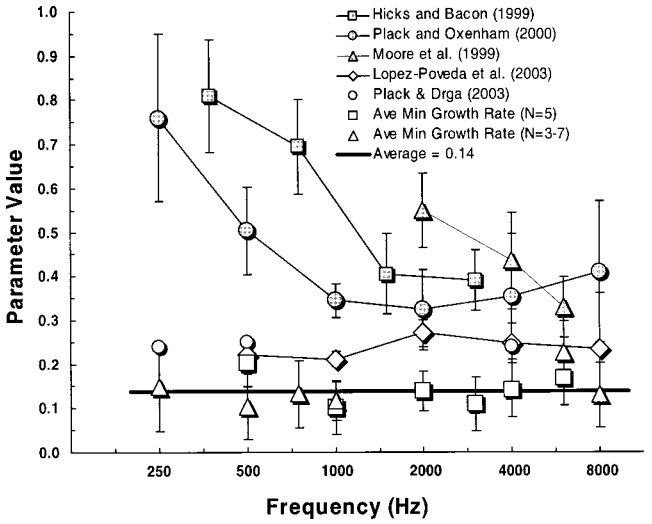


FIG. 11. The average *minimum* growth rates (maximum compression) obtained in the current experiment are shown by the unfilled squares (Group I) and the unfilled triangles (Group II) as a function of probe frequency. Error bars are standard deviations about the mean values. Average minimum growth rate across *all* subjects and probe frequencies is shown by the solid line. Response growth rates estimated by Lopez-Poveda *et al.* (2003) and by Plack and Drga (2003) with similar TMC experiments are shown by the unfilled diamonds and circles, respectively. These investigations used off-frequency TMCs from high-frequency probes as a reference for estimating compression. Response growth rate estimates obtained from *slope ratios* between on-frequency and off-frequency nonsimultaneous-masking functions are shown by the shaded symbols. Slope ratios between on- and off-frequency growth-of-forward-masking functions reported by Hicks and Bacon (1999a) and by Moore *et al.* (1999) are shown by shaded squares and shaded triangles, respectively. Slopes of growth-of-maskability pulsation-threshold functions (masker level versus probe level) from Plack and Oxenham (2000) are shown by shaded circles.

ratios to estimate compression. The filled symbols in Fig. 11 show compression estimates obtained by previous investigations using nonsimultaneous masking (Moore *et al.*, 1999; Hicks and Bacon, 1999a; Plack and Oxenham, 2000). There are two obvious differences between those results and the current estimates. First, at all probe frequencies, the previous estimates (shaded symbols) are larger than the current estimates (unfilled symbols), i.e., the previous methods underestimated the strength of compression. Second, the previous estimates become larger as probe frequency is reduced below 1000 Hz, i.e., the previous estimates indicated that compression is less strong at lower probe frequency regions. The first effect is probably due to off-frequency listening. The second is probably due to an invalid assumption about linear response growth for off-frequency maskers at low probe frequencies. Our reasoning for these conclusions follows.

A. Off-frequency listening effects

Previous estimates of compression shown in Fig. 11 involved the acquisition of growth-of-masking (GOM) functions for nonsimultaneous masker and probe tones. Two of the investigations involved forward masking (Moore *et al.*, 1999; Hicks and Bacon, 1999a), and one involved the pulsation threshold (Plack and Oxenham, 2000). These estimates of compression were derived from GOM slope ratios for on-frequency and off-frequency nonsimultaneous maskers, which were presented in a quiet background. The rationale for using this GOM slope ratio to estimate compression came from an experiment by Oxenham and Plack (1997), in one subject, in which they found that the GOM slope ratio in quiet was similar to the GOM slope ratio obtained in a background noise that masked off-frequency listening (Patterson and Nimmo-Smith, 1980; O'Loughlin and Moore, 1981; Yacullo and Abbas, 1986). Presumably off-frequency listening that affected the on-frequency GOM slope had a similar effect on the off-frequency GOM slope. Thus, the ratio between on-frequency and off-frequency GOM slopes evolved as a psychophysical estimate of compression.

For these previous estimates of compression, both the masker level and probe level covaried in a quiet background; either the probe level at masked threshold increased with masker level (Hicks and Bacon, 1999a) or the masker level at masked threshold increased with probe level (Moore *et al.*, 1999; Plack and Oxenham, 2000). When both masker and probe level co-vary, off-frequency listening can dramatically affect GOM slopes (Bacon *et al.*, 1999; Plack and Oxenham, 2000; Nelson *et al.*, 2001). This phenomenon has been confirmed by the introduction of a high-pass noise above the probe frequency to restrict ''listening'' only to frequency regions near the probe frequency (Oxenham and Plack, 1997; Plack and Oxenham, 2000; Nelson *et al.*, 2001). Specifically, Nelson *et al.* found that GOM slopes increased in the presence of a high-pass noise, more so for an off-frequency masker than for an on-frequency masker. Compression estimates were based on the slope ratios between on-frequency and off-frequency GOM functions, as was the case with the previous compression estimates noted above (Moore *et al.*, 1999; Hicks and Bacon, 1999a; Plack and Oxenham, 2000). Compression estimates were underestimated (slope ratios were larger) by about a factor of 2.3, when GOM functions were obtained in a quiet background that allowed off-frequency listening to take place (slope ratios of 0.26 and 0.35 from two subjects), than when GOM functions were obtained in a high-pass noise that restricted off-frequency

TABLE I. Average minimum response growth rates (and standard deviations) at different probe frequencies for two groups of normal-hearing listeners. Group I includes five subjects tested at each of six probe frequencies. Group II includes nine subjects tested at one or more probe frequencies.

	Probe frequency								
	250	500	750	1000	2000	3000	4000	6000	8000
Group I ($N=5$)		0.20		0.10	0.14	0.11	0.14	0.17	
Stdev		0.055		0.060	0.046	0.061	0.061	0.060	
Group II	0.15	0.10	0.13	0.12				0.22	0.13
Stdev	0.101	0.074	0.076	0.044				0.072	0.073
$N=$	5	7	7	6				5	3

listening (slope ratios of 0.11 and 0.16) (Nelson *et al.*, 2001).

It is notable that the previous estimates of compression at 1000 Hz and above, in Fig. 11, differ from the current estimates of compression, by a factor of approximately 2.5. Those previous estimates involved GOM functions in which both the masker and probe covaried, without any background noise to restrict off-frequency listening. The current estimates of compression were obtained with a technique that does not allow the probe level to vary, and thus is not influenced substantially by off-frequency listening. This was previously confirmed by Nelson *et al.* (2001) with TMCs obtained in the presence of a high-pass noise. These comparisons strongly suggest that the underestimation of compression by a factor of about 2.5 by the previous investigations was very likely due to off-frequency listening.

B. Linearity assumptions for off-frequency maskers

Regardless of issues relating to off-frequency listening, the GOM slope ratio estimates of compression are strongly dependent on the off-frequency GOM slope. For example, if compression affects the off-frequency masker response growth, at the on-frequency place, then the GOM slope for the off-frequency masker will be more gradual, and the subsequent on- versus off-frequency GOM slope-ratio will be larger, which will be interpreted as less compression. This is essentially the result exhibited in Fig. 11 by the previous estimates of compression across the probe frequency region. In two of the previous studies (Hicks and Bacon, 1999a; Plack and Oxenham, 2000), response growth rates increased with decreasing probe frequency for probe frequencies below 1000 Hz; in the other study (Moore *et al.*, 1999), response growth rates decreased progressively as probe frequency decreased from 6000 to 2000 Hz. These results are consistent with the idea proposed by Plack and Drga (2003) and Lopez-Poveda *et al.* (2003) that the masker response at the probe frequency, for a fixed Fm/Fp ratio, is less linear at lower frequency regions. Thus, estimates of compression taken from on- versus off-frequency GOM slope ratios appear to be less at low frequencies.

By way of contrast, the estimates of compression shown by unfilled symbols in Fig. 11 did not change significantly with the probe frequency region. Compression estimates from the current study are shown by the unfilled squares and triangles. Similar data from Lopez Poveda *et al.* (2003) and Plack and Drga (2003) are shown by the unfilled diamonds and circles, respectively. The similarity was largely because neither estimate of compression depended upon the assumption of linear response growth for the off-frequency maskers at lower probe frequencies.

The average compression estimates obtained by Lopez-Poveda *et al.* (2003) and Plack and Drga (2003) ranged between 0.21 and 0.27, while the estimates from the current study ranged between 0.10 and 0.22 (average$=0.14$). Their compression estimates were independent of probe frequency, as in the current study, but the compression estimates were less strong (steeper response growth rates) than the current estimates. The reason for this difference is most likely found in slight differences in the procedures used to estimate response growth rates from derived I/O curves. For example, Lopez-Poveda *et al.* fitted their derived I/O curves with up to three linear segments. The slope of a mid-range segment was used to estimate compression. The number of input levels included in that mid-range segment varied among subjects, but often included a wide range of input levels. By contrast, the current procedure fitted derived I/O curves with a third-order polynomial, which presumes that compression changes continuously with input level, and allows one to calculate compression continuously along the derived I/O curve. Specification of the *minimum* response growth rate with this procedure leads to more gradual slope estimates, largely because it does allow specification of local response growth rates. Had the current study averaged local compression over a range of input levels, the amount of compression would have been less (larger response growth rates) and more in line with estimates obtained by Lopez-Poveda *et al.* and Plack and Drga (2003). Regardless, all three studies found that their compression estimates were independent of the probe frequency region.[2]

V. SUMMARY AND CONCLUSIONS

Response growth rates obtained with isoresponse TMCs vary continuously with the input level from near linear at low levels, to highly compressive at mid-levels, to near linear again at high levels. Derived I/O curves are similar across probe frequency regions from 250 to 8000 Hz, as long as the average off-frequency TMC for high probe frequencies (2000–4000 Hz) is used to derive those growth rates. The minimum response growth rate (maximum compression) tends to remain the same at all probe frequencies, but the range of input levels with strong compression is slightly smaller at lower and higher frequencies. Compression estimates obtained by some investigators, using slope ratios between on-frequency and off-frequency nonsimultaneous GOM functions, underestimated maximum compression. In addition, those studies also underestimated compression at lower-frequency regions, largely because they relied on the assumption of linear response growth at the probe frequency for off-frequency maskers.

ACKNOWLEDGMENTS

This work was supported largely by Grant No. DC00149 from NIDCD. This research was also supported in part by the Lion's 5M International Hearing Foundation. The authors appreciate the excellent suggestions to this manuscript provided by Sid Bacon and Chris Plack.

[1]The TMC procedure for estimating compression was first introduced at a special session honoring the contributions of Robert C. Bilger: ''Bilger and Better Science'' [Nelson, D.A. and Schroder, A.C. (**1999**). ''Forward masking recovery and peripheral compression in normal-hearing and cochlear-impaired ears,'' J. Acoust. Soc. Am. **106**, 2176(A)]. That presentation acknowledged Bob's strong influence on current research in the field, particularly on the search for measures of nonlinearity in normal and impaired ears.

[2]As pointed out by previous investigators (Plack and Drga (2003) and Plack and O'Hanlon (2003)), compression at low-frequency regions indicated by psychophysical measurements could well have a central origin.

ANSI (**1996**). *S3.6-1996 Specification for Audiometers* (American National Standards Institute, New York).

Bacon, S. P., Boden, L. N., Lee, J., and Repovsch, J. L. (**1999**). ''Growth of simultaneous masking for fm<fs: Effects of overall frequency and level,'' J. Acoust. Soc. Am. **106**, 341–350.
Cooper, N. P., and Rhode, W. S. (**1995**). ''Nonlinear mechanics at the apex of the guinea-pig cochlea,'' Hear. Res. **82**, 225–243.
Cooper, N. P., and Yates, G. K. (**1994**). ''Nonlinear input–output functions derived from the responses of guinea-pig cochlear nerve fibers: Variations with characteristic frequency,'' Hear. Res. **78**, 221–234.
Hicks, M. L., and Bacon, S. P. (**1999a**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.
Hicks, M. L., and Bacon, S. P. (**1999b**). ''Effects of aspirin on psychophysical measures of frequency selectivity, two-tone suppression, and growth of masking,'' J. Acoust. Soc. Am. **106**, 1436–1451.
Khanna, S. M., and Hao, L. F., (**1999**). ''Nonlinearity in the apical turn of living guinea pig cochlea,'' Hear. Res. **135**, 89–104.
Levitt, H. (**1971**). ''Transformed up-down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (**2003**). ''Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing,'' J. Acoust. Soc. Am. **113**, 951–960.
Moore, B. C. J., and Oxenham, A. J. (**1998**). ''Psychoacoustic consequences of compression in the peripheral auditory system,'' Psychol. Rev. **105**, 108–124.
Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (**1999**). ''Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism,'' J. Acoust. Soc. Am. **106**, 2761–2778.
Murugasu, E., and Russell, I. J. (**1995**). ''Salicylae ototoxicity: The effects on basilar membrane displacement, cochlear microphonics, and neural responses in the basal turn of the guinea pig cochlea,'' Aud. Neurosci. **1**, 139–150.
Nelson, D. A., and Freyman, R. L. (**1987**). ''Temporal resolution in sensorineural hearing-impaired listeners,'' J. Acoust. Soc. Am. **81**, 709–720.
Nelson, D. A., and Schroder, A. C., (**1997**). ''Linearized response growth inferred from growth-of-masking slopes in ears with cochlear hearing loss,'' J. Acoust. Soc. Am. **101**, 2186–2201.
Nelson, D. A., Schroder, A. C., and Wojtczak, M. (**2001**). ''A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **110**, 2045–2064.
O'Loughlin, B. J., and Moore, B. C. J. (**1981**). ''Off-frequency listening: Effects on psychoacoustical tuning curves obtained in simultaneous and forward masking,'' J. Acoust. Soc. Am. **69**, 1119–1125.
Oxenham, A. J., and Plack, C. J. (**1997**). ''A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **101**, 3666–3675.
Patterson, R. D., and Nimmo-Smith, I. (**1980**). ''Off-frequency listening and auditory-filter asymmetry,'' J. Acoust. Soc. Am. **67**, 229–245.
Plack, C. J., and Drga, V. (**2003**). ''Psychophysical evidence for auditory compression at low characteristic frequencies,'' J. Acoust. Soc. Am. **113**, 1574–1586.
Plack, C. J., and O'Hanlon, C. G. (**2003**). ''Forward masking additivity and auditory compression at low and high frequencies,'' J. Assoc. Res. Otolaryngol. **4**, 405–415.
Plack, C. J., and Oxenham, A. J. (**2000**). ''Basilar-membrane nonlinearity estimated by pulsation threshold,'' J. Acoust. Soc. Am. **107**, 501–507.
Rhode, W. S., and Cooper, N. P. (**1996**). ''Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*,'' Aud. Neurosci. **3**, 101–121.
Rhode, W. S., and Recio, A. (**2000**). ''Study of mechanical motions in the basal region of the chinchilla cochlea,'' J. Acoust. Soc. Am. **107**, 3317–3332.
Rhode, W. S., and Robles, L. (**1974**). ''Evidence from Mossbauer experiments for nonlinear vibration in the cochlea,'' J. Acoust. Soc. Am. **55**, 588–596.
Robles, L., Ruggero, M. A., and Rich, N. C. (**1986**). ''Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases,'' J. Acoust. Soc. Am. **80**, 1364–1374.
Ruggero, M. A. (**1992**). ''Responses to sound of the basilar membrane of the mammalian cochlea,'' Curr. Opin. Neurobiol. **2**, 449–456.
Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.
Russell, I. J., and Nilsen, K. E. (**1997**). ''The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane,'' Proc. Natl. Acad. Sci. U.S.A. **94**, 2660–2664.
Yacullo, W. S., and Abbas, P. J. (**1986**). ''Detection cues in forward masking and their relationship to off-frequency listening,'' J. Acoust. Soc. Am. **80**, 452–465.
Yates, G. K. (**1990**). ''Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions,'' Hear. 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,'' Hear. Res. **45**, 203–219.
Zinn, C., Maier, H., Zenner, H.-P., and Gummer, A. W. (**2000**). ''Evidence for active, nonlinear, negative feedback in the vibration response of the apical region of the *in-vivo* guinea-pig cochlea,'' Hear. Res. **142**, 159–183.

The temporal effect with notched-noise maskers: Analysis in terms of input–output functions[a)]

Elizabeth A. Strickland[b)]
Department of Audiology and Speech Sciences, Purdue University, West Lafayette, Indiana 47907-2038

(Received 5 March 2003; accepted for publication 9 February 2004)

This study examines whether a temporal masking effect may be consistent with a decrease in gain at the masker frequency during the course of the masker. Threshold level of a long-duration notched-noise masker needed to mask a 1- or 4-kHz signal was measured for three conditions: a short-duration signal with a short delay or a long delay from masker onset, and a long-duration signal. The difference between threshold for the long-delay signal and the short-delay signal was defined as the temporal effect. The size of the temporal effect depended on signal frequency, signal level, and masker notch width. Filters estimated from the data had narrower bandwidths for the long-delay condition than for the short-delay condition or the long-duration condition, which seems inconsistent with the hypothesis of a decrease in gain. However, modeling of the data in terms of basilar-membrane input–output functions is consistent with a decrease in gain in the masker frequency region during the course of the masker. For a notch width of 0.0 the results are consistent with a decrease in gain at the signal frequency. For a relative notch width of 0.4, the decrease in gain at the masker frequency may cause a decrease in the suppression of the signal. This decrease in suppression could explain the decrease in filter bandwidth with signal delay. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1691036]

PACS numbers: 43.66.Dc, 43.66.Mk, 43.66.Ba [MRL] Pages: 2234–2245

I. INTRODUCTION

As a brief signal is delayed from the onset of a longer-duration masker, the threshold signal-to-masker ratio may decrease (e.g., Scholl, 1962). This change in threshold has been called overshoot (Zwicker, 1965) or the temporal effect (Wright, 1995b). The term temporal effect will be used in this paper. If the masker is a notched noise and the signal is a tone, the temporal effect generally increases as the notch width increases (Hicks and Bacon, 1992; Bacon *et al.*, 2002) although for 1-kHz signals some studies have shown an increase and then a decrease (Carlyon, 1989) or no change (Moore *et al.*, 1987). If the masker and the signal are both tones, this effect is seen mainly when the masker is higher in frequency than the signal, and again increases as the masker frequency is moved away from the signal frequency (Bacon and Viemeister, 1985; Bacon and Moore, 1986; Kimberley *et al.*, 1989; Bacon *et al.*, 2002). These results are quite interesting because they suggest that something is changing in the auditory system as it is being stimulated by sound, and that this change affects psychophysical measures of frequency selectivity.

Although the temporal effect may well be due to processes at more than one level of the auditory system, there is evidence that it is related to the active process in the cochlea. This active process is thought to be due to amplification by the outer hair cells, and is responsible for high sensitivity and frequency selectivity. For listeners with permanent cochlear hearing loss the temporal effect is decreased or abolished, when measured for a tone in a broadband noise (Bacon and Takahashi, 1992) or for tone-on-tone masking (Bacon *et al.*, 1988; Kimberley *et al.*, 1989). The temporal effect for a tone in a broadband noise may also be temporarily decreased or abolished by ingestion of aspirin (McFadden and Champlin, 1990), and by temporary threshold shift produced by exposure to an intense sound (Champlin and McFadden, 1989). All of these effects increase quiet thresholds and decrease frequency selectivity for long-duration tones, and thus are thought to turn down the amplification due to the active process. When the temporal effect is decreased, it is because thresholds improve for signals at masker *onset*. Therefore, it has been hypothesized that amplification due to the active process is high at masker onset, and decreases during the course of the masker (von Klitzing and Kohlrausch, 1994; Strickland, 2001b). This could be mediated by the medial olivocochlear bundle (MOCB) of the efferent system, which will be discussed later.

Although this hypothesis is appealing, there is an apparent discrepancy in the results noted above. The temporal effect tends to grow with masker notch width; thus, psychophysical frequency selectivity is *sharper* when a signal is delayed from the onset of a tonal or notched-noise masker than it is when the signal is at masker onset. If psychophysical frequency selectivity is taken as a measure of gain at the signal frequency, it would seem that it increases with delay of the signal from masker onset, rather than decreases.

One problem with this interpretation is that the masker would be turning down gain in the *masker* frequency region, and this frequency region changes as the masker notch width changes. This was addressed in a previous study (Strickland, 2001b), in which the temporal effect was measured for just one condition, with no notch in a broadband masking noise.

[a)]Portions of this research were presented at the 139th Meeting of the Acoustical Society of America, Atlanta, GA.
[b)]Electronic mail: estrick@purdue.edu

The notched-noise method was then used to measure thresholds with and without a fixed-level broadband noise precursor. Frequency selectivity was broader following the precursor than with no precursor, which supported the hypothesis that gain at the signal frequency was decreased during the course of the precursor.

A second problem with the above interpretation is that psychophysical frequency selectivity may not solely depend on gain at the *signal* frequency. When the stimulation before the signal (the masker or precursor) is not at the signal frequency, it is important to recognize that in simultaneous masking, thresholds may reflect suppression as well as excitation. Physiologically, suppression refers to the fact that the response to one tone may be decreased by the addition of a second sound at a different frequency (Sachs and Kiang, 1968; Delgutte, 1990). Of particular relevance in studying changes in gain is the fact that suppression appears to depend on gain at the suppressor (in this case the masker) frequency as well as the signal frequency. Physiological data have shown that exposing the frequency region of a suppressor to kanamycin (Dallos *et al.*, 1980) or to a brief, high-intensity tone (Robertson and Johnstone, 1981) increased the suppressor level necessary to produce a given amount of suppression of a signal, without affecting the tuning curve at the signal frequency. Psychophysical data in humans also seem to be consistent with the idea that gain in the masker frequency region may play a role in the temporal effect. Bacon *et al.* (1988) measured the temporal effect for a tonal masker with a frequency of 1.2 times the signal frequency, a frequency ratio at which masking may be partially suppressive. They found that the size of the temporal effect was reduced in listeners with hearing loss in the frequency region of the masker, whether or not the loss included the signal frequency.

If amplification due to the active process decreases as the signal is delayed from masker onset, this should be apparent in other ways. Another consequence of the active process is nonlinear growth of excitation to a tone at the characteristic frequency (CF) for a place on the basilar membrane. The input–output function grows with a slope near 1 for low-input levels, grows more slowly for mid-input levels, and then may grow more rapidly again for high-input levels (e.g., Ruggero *et al.*, 1997). Several studies have derived input–output functions from psychophysical data by measuring growth of masking functions for forward masking using on-frequency and off-frequency maskers (Oxenham and Plack, 1997; Nelson *et al.*, 2001). Nonlinear input–output functions for signals at CF have been used fairly successfully to model data from simultaneous as well as forward masking (Moore *et al.*, 1996; Glasberg and Moore, 2000; Wojtczak *et al.*, 2001). In the Strickland (2001b) study discussed above, derived input–output functions were less compressive following the precursor than with no precursor, consistent with the idea that the broadband precursor decreased gain at the signal frequency.

The purpose of the present study was to examine the hypothesis that the temporal effect is partly due to a decrease in amplification at the masker frequency during the course of the masker, even when the stimulation preceding the signal has a notch in the frequency spectrum centered at the signal frequency. The notched-noise method was used to study the temporal effect as a function of signal frequency, signal level, and notch width. By using multiple signal levels, it should be possible to analyze the results in terms of input–output functions. These may be useful in quantifying whether the results are consistent with a decrease in suppression when there is a notch in the masker. This study is similar to one done by Bacon and Viemeister (1985). That study, however, used only a 1-kHz signal, and tonal maskers. In the present study, signal frequencies of 1 and 4 kHz will be used, because measures of the active process tend to show larger effects at higher frequencies. Also, the masker will be a notched noise, because this limits off-frequency listening, and other problems such as detection of combination tones and beats. Other studies have used only one fixed level for the masker (Moore *et al.*, 1987; Carlyon, 1989; Hicks and Bacon, 1992; Wright and Dai, 1994) or signal (Bacon and Moore, 1986; Kimberley *et al.*, 1989). Therefore, the purpose of the present study was to examine the temporal effect in more detail, by fixing the signal level across a range of levels, and varying the level of a notched-noise masker to find threshold. The data will also be analyzed in terms of filter shapes, to determine the uses and the limitations of this method.

II. METHODS

A. Listeners

Seven listeners participated in the experiment. Most listeners participated in only a subset of the conditions. All listeners had thresholds within laboratory norms for 200-ms pure tones at octave frequencies from 250 to 8000 Hz in the ear tested, which was the left ear. The age range was 19 to 42 years, with a median of 22 years. All listeners except for L4, who is the author, were paid for their participation.

B. Stimuli

The signal was a sinusoid with a frequency of 1 or 4 kHz. The masker was a noise with a spectral notch around the signal frequency. The outer spectral edges of the noise were fixed at $0.2*f_s$ and $1.8*f_s$, where f_s is the signal frequency. The spectral edges of the notch closest to the signal are expressed in normalized units of $|f-f_s|/f_s$, where f is the edge of the notch. This normalized unit will be referred to as Δf. Notch edges were placed symmetrically around the signal, and were set at Δf values of 0.0, 0.1, 0.2, 0.3, and 0.4. The average long-term spectrum of the masker was flat within the passbands.

Masker thresholds were measured for three temporal relations between the signal and masker, as schematized in Fig. 1. All masker and signal durations include 5-ms $\cos^2$ onset and offset ramps, and are measured between 0-voltage points. In order to measure the temporal effect, a 10-ms signal was used. This duration was chosen to be brief, yet long enough to avoid effects of spectral splatter (Bacon and Viemeister, 1985). In the short-duration, long-delay condition the masker duration was 400 ms. The signal onset was delayed 202 ms from the masker onset. In the short-duration,

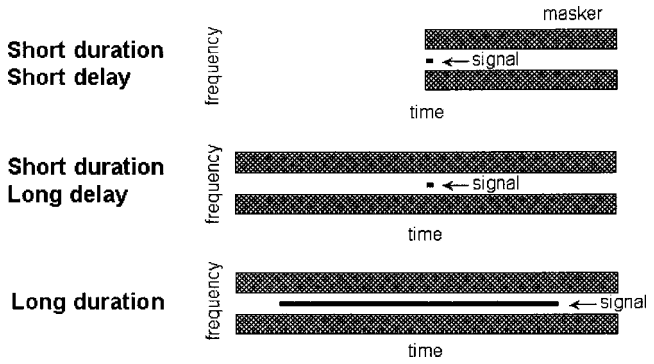


FIG. 1. Schematic showing the stimulus configuration in the three conditions.

short-delay condition, the masker duration was 200 ms, and the signal onset was delayed 2 ms from the masker onset. Thus, in the short- and long-delay conditions, the delay between the signal and the masker offset was kept constant, in order to avoid any offset effects (e.g., Wright and Dai, 1994). Signal level was set at 40, 50, 60, and 70 dB SPL. Thresholds were also measured for a long-duration signal in a long-duration masker. Although a great deal of data is available for this condition, it has often been measured with masker level fixed, while in the present experiment the signal level was fixed. Also, because there is some individual variability, it was desirable to be able to compare data across conditions for the same listeners. In the long-duration condition, two different signal and masker durations were used, due to a programming error. For L1, L2, and L3, the signal duration was 350 ms, and the masker duration was 655 ms. For L4, the signal duration was 300 ms and the masker duration was 400 ms. In both cases, the signal onset was delayed from the masker onset by 50 ms. This small difference in signal duration should not have a significant effect on thresholds. Signal level was fixed at 20, 30, and 40 dB SPL.

The signal and masker were digitally produced, and were output through two separate D/A channels (TDT DA3-4). They were low-pass filtered at 10 kHz (TDT FT5). The level was controlled by programmable attenuators (TDT PA4). In addition, for low signal levels, additional attenuation of the masker was provided by a manual attenuator (Leader LAT-45). The signal and masker were mixed (TDT SM3) and led to a headphone buffer (TDT HB6). The stimuli were presented to one of two ER-2 insert earphones. These earphones have a flat frequency response from 250 to 8000 Hz.

C. Procedure

Listeners were seated in a double-walled sound-attenuating booth. An adaptive three-interval forced-choice procedure with a two-up, one-down stepping rule was used to measure thresholds, which converges on the 70.7%-correct point of the psychometric function (Levitt, 1971). Temporal intervals were marked visually on a computer monitor, and subjects responded via a computer keyboard. Visual feedback was provided. The initial masker level was set below estimated threshold. The initial step size was 5 dB, and decreased to 2 dB after the first four reversals. Each run consisted of 50 trials. Threshold was taken as the mean of the last even number of reversals at the smaller step size. If the standard deviation of the run was greater than 5 dB, it was not included in the threshold estimate. Thresholds were estimated from three or four runs.

For L1, L2, and L3, the condition being measured was usually constant within a session, but was randomized across sessions. In general, one signal frequency and level were chosen, and the notch widths were measured in random order for that signal. For L4, L5, L6, and L7, the strategy was different. These listeners were run within a limited time span, so data were collected so as to have a complete set within a condition before moving to the next condition. Also, in order to make the task more consistent for the listeners, generally the notch width was fixed within a session. The signal level was usually increased from low to high across the session. This was done to avoid any effects of exposure to a higher masker level affecting the responses to a lower masker level (Strickland, 2001b). Due to a programming error, for L1, L2, and L3, in the short-duration conditions the masker was gated rectangularly rather than with a $\cos^2$ ramp. Their long-delay data are similar to the data of the other listeners, but the short-delay data are not. L1 was rerun with the proper gating in the short-delay condition, and only those results are included here. Data for the short-delay condition from L2 and L3 were not used.

III. RESULTS AND DISCUSSION

A. Data

In Fig. 2, results for the short-duration 1-kHz signal are shown for the four listeners who were run in both the short-delay and the long-delay conditions. The short-delay results are shown by the open symbols, and the long-delay results by the filled symbols. Only two signal levels are shown, for clarity; 40 dB SPL (circles) and 60 dB SPL (downward triangles). L6 is missing two data points because a consistent threshold could not be obtained in those conditions. The temporal effect is defined as the difference between the long-delay and short-delay thresholds. The amount of temporal effect ranges from nearly none (L4) to more than 20 dB (L6), and is consistent with the variability seen in previous studies (Carlyon, 1989; Hicks and Bacon, 1992). When a temporal effect is seen, it tends to be larger when the signal level is 60 dB SPL than when it is 40 dB SPL, and tends to remain constant for $\Delta f = 0.1$ and greater.

In Fig. 3, results are shown for the 4-kHz signal, for all listeners who were run in both the short-delay and long-delay conditions. Symbols and levels are the same as those in Fig. 2. At this frequency, all listeners showed a temporal effect for both signal levels. When the signal level was 40 dB SPL, the temporal effect increased with Δf. The temporal effect increased with level when $\Delta f = 0.0$, but decreased with level when $\Delta f = 0.3$ and 0.4. As in this study, Hicks and Bacon (1992) found that all of their listeners showed a temporal effect for a 4-kHz signal.

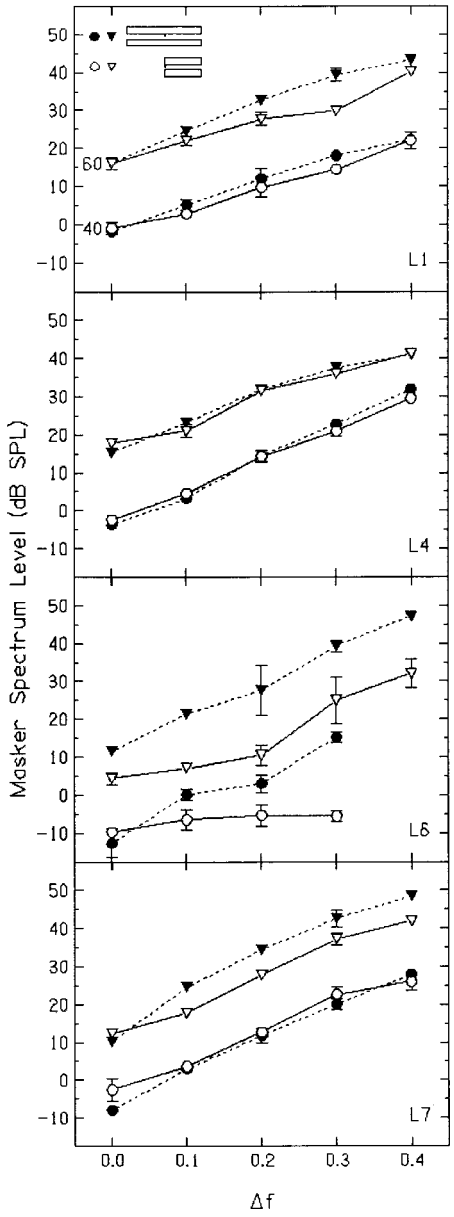


FIG. 2. Threshold masker level for detection of a 1-kHz short-duration signal in the short-delay (open symbols) and long-delay (filled symbols) conditions. Results are shown for a signal level of 40 (circles) and 60 (downward triangles) dB SPL.

B. Filter estimates

Filter functions were estimated from the data as a way of quantifying the results, and also so that they could be compared to previous results. Functions were estimated with an adaptive fitting program called POLYFIT (Rosen *et al.*, 1998), using the assumption that filter shape is a function of signal level. The basic function used was a rounded exponential, as shown in Eq. (1)

$$W(g)=(1-w)(1+(pg))e^{-pg}+w(1+tg)e^{-tg}, \qquad (1)$$

where g is frequency normalized to the center frequency, p is for the slope of the tip of the filter, t for the tail of the filter, and w determines where the filter becomes more dominated by the tail than the tip (Patterson *et al.*, 1982). Another parameter used in fitting was k, which is the threshold signal-to-masker ratio in dB. Data for all signal levels for a given

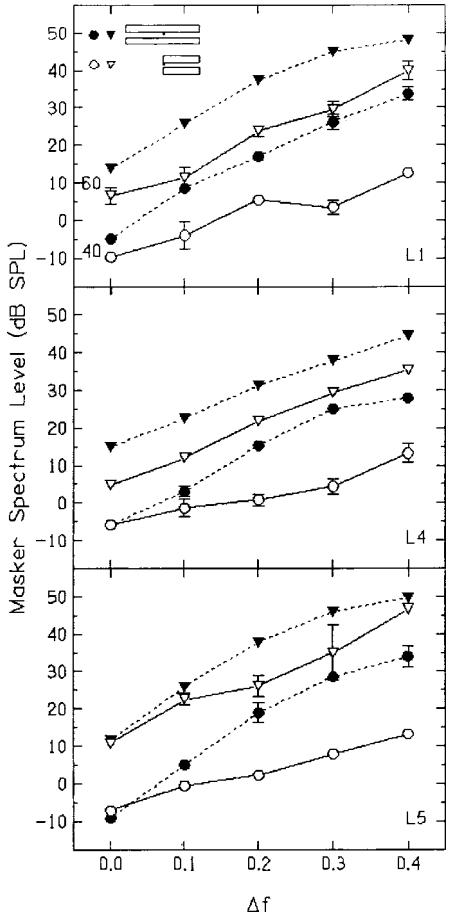


FIG. 3. As in Fig. 2, for a 4-kHz signal.

listener and condition were fit together. Each term in the equation above was a polynomial dependent on signal level that could be adjusted up to degree three. Each polynomial term was adjusted to minimize the rms error. The exception to this was t, which was fixed at a single value with level. The equivalent rectangular bandwidth (ERB) and k from the best-fitting model for individual and mean data for the short-duration conditions are shown in Table I. For a given signal level, the ERB tends to be smaller in the long-delay condition than in the short-delay condition, even for listeners who showed only very small temporal effects. The rms error is also consistently lower for the long-delay condition than for the short-delay condition. For most listeners, the ERB tended to decrease with signal level. This was also found by Bacon and Viemeister (1985). Q_{ERB} values (signal frequency/ERB) calculated from the data in Table I are similar to Q_{ERB} values calculated from data from Wright and Dai (1994) for a 2.5-kHz signal under comparable conditions (their 5-ms signal, short-delay masker and long-delay masker conditions) and values calculated from data of Bacon *et al.* (2002) for a 2-kHz signal (notched-noise conditions).

Note that the values of k in Table I are much larger than those typically reported for long-duration signals. The value of k has been shown to increase as signal duration decreases, and thus it may reflect effects of temporal integration (Hant *et al.*, 1997). Also, k may depend on the input–output function at the signal frequency.

Estimated filter parameters for the long-duration signal

TABLE I. Parameters for roex filters estimated from the individual and mean short-duration data.

Frequency (kHz)	Listener	Delay	Signal level (dB SPL) 40 ERB	40 k	50 ERB	50 k	60 ERB	60 k	70 ERB	70 k	rms error
1	L1	short	237	17.6	194	19.6	176	21.6	191	23.6	1.32
		long	151	19.6	150	20.3	150	23.9	149	19.3	0.92
	L2	long			141	27.7	126	27.9			0.25
	L3	long			143	22.1	128	24.1			0.52
	L4	short	168	20.4	171	21.4	184	20.7	257	18.5	1.13
		long	151	22.4	136	24.6	123	25.5	118	25.1	0.78
	L6	short	458	24.1	223	29.6	191	34.9	198	39.9	2.51
		long	150	29.8	102	29.6	91	29.5	91	29.3	2.05
	L7	short	150	22.5	150	24.5	150	27.8	150	32.5	1.14
		long	112	28.6	105	27.3	99	28.4	93	31.9	0.85
	Mean	short	201	21.3	194	23.5	198	25.7	214	27.8	1.03
		long	155	24.0	132	25.1	120	26.3	116	27.5	0.64
4	L1	short	684	21.8	616	23.9	606	27.4	604	32.5	1.56
		long	436	18.1	412	19.3	394	20.5	385	21.7	0.91
	L2	long			546	18.9	475	21.3	424	23.3	1.19
	L3	long	574	15.0	420	17.8	337	20.7			0.82
	L4	short	1206	15.3	779	20.6	727	25.8	685	31.1	1.57
		long	491	19.5	495	18.3	504	19.1	524	22.0	0.92
	L5	short	1076	16.1	806	18.2	661	20.2	623	22.3	1.70
		long	377	23.1	377	23.1	377	23.1	380	23.1	0.75
	Mean	short	1128	16.2	771	20.5	649	24.7	619	28.9	0.73
		long	483	18.6	462	19.1	444	20.3	431	22.4	0.64

are shown in Table II. Note that ERBs are larger on average for the long-duration condition than for the short-duration, long-delay condition. This was also observed by Wright and Dai (1994).

The filter estimates show that bandwidths are smaller in the long-delay condition than the short-delay or the long-duration condition. It is clear, however, that they do not directly reflect the excitation produced by a single tone in the cochlea, because bandwidths also decrease with level, while physiological responses to single tones broaden with signal level (e.g., Ruggero *et al.*, 1997). The filter estimates also do not seem to fit with other evidence that suggests that amplification due to the active process may be decreased in the long-delay condition. The data must be analyzed in a different manner to explore this hypothesis.

TABLE II. Parameters for roex filters estimated from the individual and mean long-duration data using Eq. (1).

Frequency (kHz)	Listener	rms error	Signal level (dB SPL) 20 ERB	30 ERB	40 ERB
1	L1	0.591	185	191	198
	L2	0.366	136	150	169
	L3	0.184		162	163
	L4	0.660	125	131	153
	Mean	0.172	146	159	169
4	L1	0.816	484	468	481
	L2	0.730	766	580	527
	L3	0.494		508	528
	L4	0.894	659	509	611
	Mean	0.644	647	517	540

C. Input–output functions

1. $\Delta f = 0.0$

Measurements were made at multiple signal levels because this allows for interpretation in terms of basilar-membrane input–output functions. It is already known that this approach produces results consistent with a decrease in gain at the signal frequency during the course of the masker when $\Delta f = 0.0$, as was suggested by von Klitzing and Kohlrausch (1994) and quantitatively modeled by Strickland (2001b). The approach taken will be to use this condition to quantify changes in gain when the signal and masker are at the same frequency, then determine whether the same changes in gain would apply at the masker frequency when $\Delta f = 0.4$. Results for the 4-kHz signal will be examined first, because the temporal effect is larger than for the 1-kHz signal. When $\Delta f = 0.0$, it will be assumed that masking is excitatory, although in reality suppression may also play a role.

In Fig. 4, results are shown when $\Delta f = 0.0$ for a 4-kHz signal as a function of signal level. Included are results for the three listeners in Fig. 3 as well as L6 and L7, who did not have complete results for the larger notch widths. The open circles show the short-delay condition, and the filled circles the long-delay condition. The dashed lines are model estimates which will be discussed later. At mid- to high-signal (and masker) levels, the masker level at threshold is lower in the short-delay condition than in the long-delay condition. There is a great deal of individual variability in the short-delay thresholds, and less in the long-delay thresholds. These results are consistent with previous data (Bacon, 1990; von Klitzing and Kohlrausch, 1994; Strickland, 2001b). For four of the five listeners (all but L5) there are subtle slope

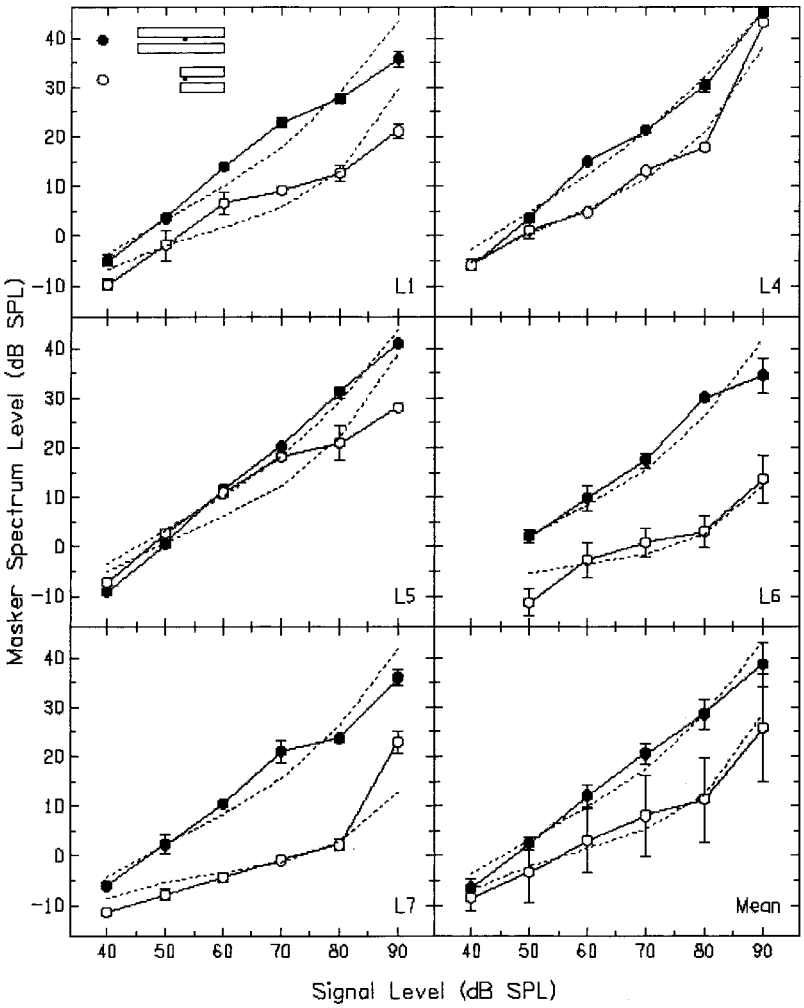


FIG. 4. Threshold masker level for detection of a 4-kHz short-duration signal as a function of signal level when $\Delta f = 0.0$. Results for the short-delay condition are shown by the open circles, and those for the long-delay condition by the filled circles. The dashed lines are model predictions.

changes even in the long-delay condition. This was also found by von Klitzing and Kohlrausch (1994), who showed a gradual approach to linearity in the function as the signal was delayed from masker onset.

As was suggested by von Klitzing and Kohlrausch (1994), and shown in a previous paper (Strickland, 2001b), data for both the short-delay and long-delay conditions for $\Delta f = 0.0$ may be fairly well approximated by using a model incorporating a psychophysical estimate of the basilar-membrane input–output function for a tone at the characteristic frequency (CF). It is assumed that thresholds are based on some constant internal difference between the response to the signal plus masker and the masker alone. This means there must be some transformation of the stimuli within the auditory system, since this difference is not constant for the physical stimuli.

Such a transformation occurs in the cochlea. Physiological data show that amplification in the cochlea is greatest at low levels and decreases with level (e.g., Ruggero *et al.*, 1997). A function of this form will be estimated by assuming that the signal-plus-masker to masker ratio at the output is constant. In the present experiment, assume that when $\Delta f = 0.0$, the energy that dominates in masking the signal is at and near the signal frequency. The signal plus masker and the masker can then be assumed to follow the same input–output function at the signal place. For the short-duration signals, the signal-plus-masker level at threshold is much higher than the masker level; therefore, the signal level will be used as the estimate of the signal-plus-masker input level. The level of the masker was estimated to be the level passing through an ERB centered at the signal frequency. The ERB estimate was 517 Hz, the ERB shown in Table II for a long-duration signal at 30 dB SPL.[1] The goal is to find an input–output function which produces a constant difference between the signal plus masker and masker levels at the output, given the signal and masker inputs shown in Fig. 4.

A modification of the function used by Glasberg and Moore (2000) was found to fit the data well. The gain formula used was

$$G_{\mathrm{dB}} = -0.1L + A + B(1 - (1/(1 + \exp(0.05(60 - L))))), \tag{2}$$

where L is the input level in dB, and A and B depend on the maximum gain, $G_{\max}$, in the following way so that G_{dB} is equal to $G_{\max}$ when $L = 0$, and zero when $L = 100$:

$$A = -0.1430 G_{\max} + 11.430, \tag{3}$$

$$B = 1.1999 G_{\max} - 11.999. \tag{4}$$

The total output is equal to L plus G_{dB}. The slope of the function is near 0.9 at low and high levels. The central position of the shallower slope of the function is determined by

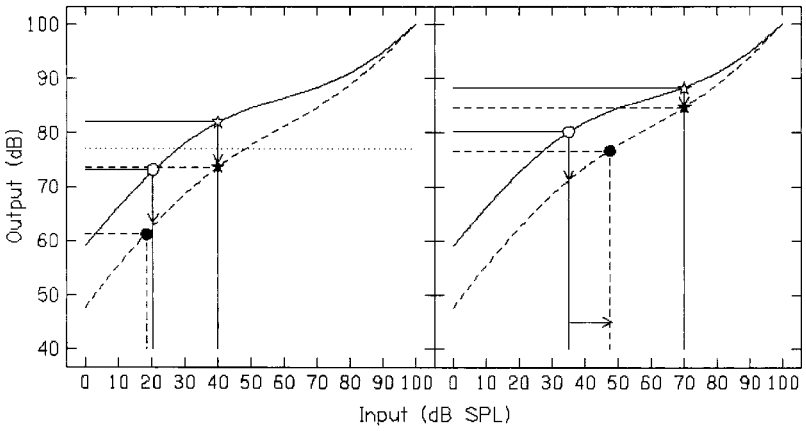


FIG. 5. Schematic illustrating how the long-delay data for $\Delta f = 0.0$ may be fit with an input–output function with a lower G_{max} than that for the short-delay condition. Only two of the six signal levels are shown, for clarity. The signal is shown by stars, the masker by circles, open and filled symbols are for the short-delay and long-delay condition, respectively. G_{max} is 59 dB for the solid input–output function, and 47 dB for the dashed input–output function. The signal level is 40 dB SPL in the left panel, and 70 dB SPL in the right panel. The dotted line in the left panel is the signal output level at threshold in quiet.

the 60. This value provided a better fit to these data than the value of 50 used by Glasberg and Moore. The 0.05 term determines the extent of the region of shallower slope.

Equation (2) was used to determine how G_{max} changed across conditions, assuming that the criterion difference between the output levels of the signal plus masker (O_{S+M}) and the masker (O_M) was constant. In order to set the criterion difference, however, an initial estimate of G_{max} was needed. G_{max} was set to the value found by Glasberg and Moore (2000) in fitting notched-noise data of Baker *et al.* (1998) for long-duration signals and maskers at multiple levels.[2] This value of G_{max} was 47.7 dB. Equation (2) was then used to calculate O_{S+M}-O_M for the long-delay data from Fig. 4 (excluding the 40-dB SPL signal level, as will be discussed below). The long-delay data were chosen because in the present study the long-delay filter parameters were similar to the long-duration filter parameters. The mean value of O_{S+M}-O_M was 9.05 dB. The criterion was then fixed at this value, and for a given condition, data for all signal levels were fit with one function by allowing G_{max} to vary to minimize the standard deviation of the O_{S+M}-O_M values about the criterion.

Figure 5 illustrates this fitting procedure schematically for a signal of 4 kHz. Although data for all levels were fit together, data for just two signal levels are shown as an example; 40 dB SPL in the left panel, and 70 dB SPL in the right panel. Consider first the short-delay condition in both panels, shown by the open symbols and the solid lines. The star represents the signal, and the circle the masker. The input values are calculated from the data in Fig. 4, and the output values are calculated from the input–output function. This function, shown by the curved line, is fit so that the difference between the two solid lines at the output stays constant, even though the difference between the lines at the input is approximately 20 dB in the left panel and more than 30 dB in the right panel. Now consider the long-delay condition in the right panel, shown by the solid symbols and dashed lines. G_{max} must be decreased to keep the difference between the dashed lines at output the same as the difference between the solid lines.[3] When G_{max} is turned down, O_M is decreased more than O_{S+M}, shown by the downward arrows, and the masker input level must be turned up, shown by the horizontal arrow, to maintain the same output criterion.

This would be the case at most signal levels. However, when the signal level is 40 dB SPL, in the left panel, it can be seen that the difference between the dashed lines is greater than the difference between the solid lines. The suggests that a different criterion is being used at this level. The average quiet threshold for the short-duration 4-kHz signal was 28 dB SPL. If the solid function is used as an estimate of the input–output function at the signal frequency in quiet, the output level at threshold is 77 dB, as shown by the dotted line. This is slightly above the output level for the long-delay signal (filled star), suggesting that it would also be at threshold. That is, at this low level, O_{S+M}-O_M is not being used as the criterion. The masker appears to decrease the excitation to the signal below threshold. Therefore, this condition was not used in the fitting procedure.

The values of G_{max} obtained by the fitting procedure are shown in Table III. All listeners showed a decrease in G_{max} between the short-delay and long-delay conditions, consistent with the hypothesis that the temporal effect is due to a decrease in gain at the masker frequency. In addition, the dashed lines in Fig. 4 show masker values that would produce a constant value of O_{S+M}-O_M, as predicted by Eq. (2) using the obtained value of G_{max}. The data values are predicted well in general.

The same procedure was used to fit the 1-kHz data in Fig. 6. The ERB estimate was 159 Hz (from Table II), the

TABLE III. G_{max} values for input–output functions fit to the short-duration data.

Frequency (kHz)	Listener	Delay			
		Short		Long	
		G_{max}	rms error	G_{max}	rms error
1	L1	34.74	0.33	35.81	0.73
	L4	31.28	1.51	38.24	1.32
	L6	53.87	1.08	45.48	1.68
	L7	45.44	1.10	44.25	1.38
	Mean	44.96	1.67	**43.40**	1.51
4	L1	58.56	1.45	47.23	1.73
	L4	53.71	0.76	43.47	0.77
	L5	52.95	1.83	46.93	1.04
	L6	64.85	1.62	49.76	1.24
	L7	64.68	1.19	49.80	1.36
	Mean	58.99	0.78	**47.70**	1.08

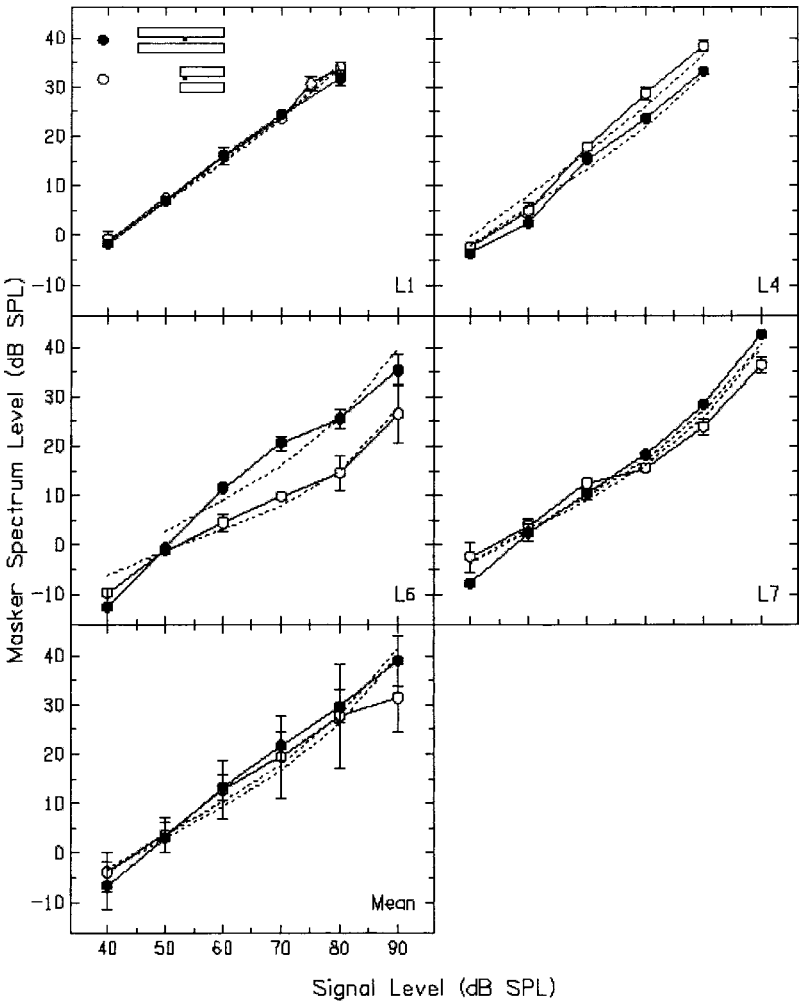


FIG. 6. As in Fig. 4, for a 1-kHz signal.

value of G_{max} from Glasberg and Moore (2000) used to determine the criterion was 43.4 dB, and the criterion value of O_{S+M}-O_M was 12.92 dB. The values of G_{max} obtained by the fitting procedure are shown in Table III, and predicted masker values are shown by the dashed lines in Fig. 6. G_{max} for the mean long-delay data for both signal frequencies is shown in bold, because those values were set by the Glasberg and Moore (2000) data. In the short-delay condition, G_{max} is 14 dB higher for the 4-kHz signal than for the 1-kHz one. All listeners show less decrease in G_{max} when the signal is 1 kHz than when it is 4 kHz (or for L1, actually show an increase for 1 kHz). This is consistent with other psychophysical (Hicks and Bacon, 1999a, b) and physiological evidence (Cooper and Yates, 1994; Cooper and Rhode, 1995; Rhode and Cooper, 1996) that effects related to the active process are more evident at higher frequencies than at lower frequencies.

2. $\Delta f = 0.4$

The results for the larger notch widths were of particular interest because they have not been examined previously in terms of input–output functions, and because suppression may play a role here. When $\Delta f = 0.4$, the masker spectrum does not have energy at the signal frequency. Previous psychophysical (Nelson *et al.*, 2001) and physiological (e.g., Ruggero *et al.*, 1997) data suggest that the response to the masker at the signal place would be linear with a slope of 1. This means that the excitatory masking will not change with masker duration before the signal. Any change in masker threshold between the short-delay and the long-delay conditions is probably due to a change in the suppression produced by the masker.

In Fig. 7, individual data are shown for the short-delay (open circles) and long-delay (filled circles) condition when $\Delta f = 0.4$. An estimate of the lowest masker level at which masking may be excitatory is shown by the dotted lines in each panel.[4] Above this masker level, masking may be excitatory and suppressive. For masker values below the line, masking is probably entirely suppressive, that is, the masker is suppressing the excitation produced by the signal so that it is at the output level for quiet threshold. This is the same effect that was shown in the long-delay condition in the left panel of Fig. 5. It is likely that the high-frequency side of the notched-noise masker is dominating in the suppression, because high-side suppressors are effective at a lower suppressor level than low-side suppressors (Cooper, 1996). In the short-delay condition, the fact that the masker level must be increased 16 dB when the signal level is increased 10 dB (from 40 to 50 dB SPL) is consistent with the slow growth rate of suppression observed by Nuttall and Dolan (1993) for suppressors at this frequency ratio above the signal frequency. The increase in masker level between the short-delay

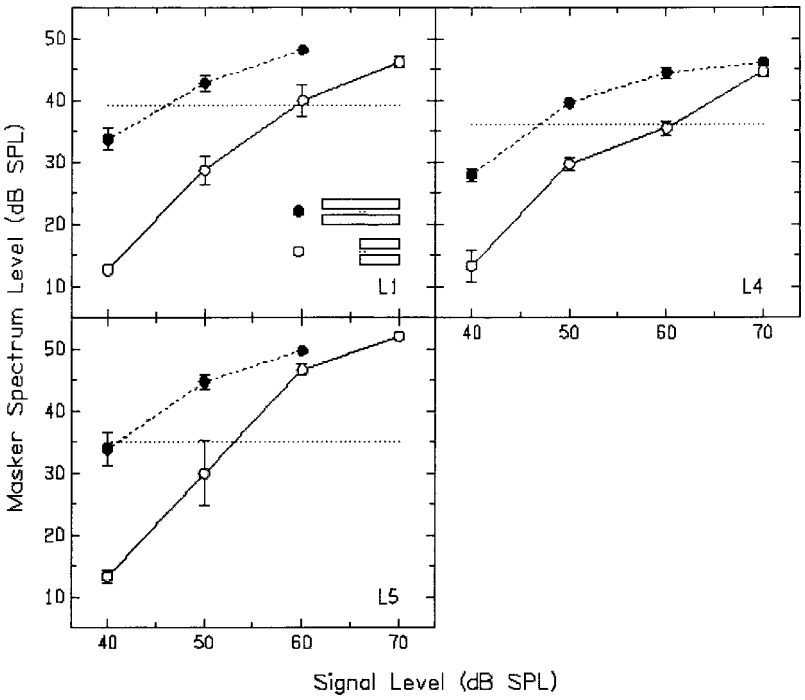


FIG. 7. Threshold masker level for detection of a 4-kHz short-duration signal as a function of signal level when $\Delta f = 0.4$. Results for the short-delay condition are shown by the open circles, and those for the long-delay condition by the filled circles. Dotted lines in each panel show estimates of the masker threshold that would be required to produce excitatory masking.

and long-delay conditions suggests that suppression decreased during the course of the masker.

A decrease in suppression in the long-delay condition would be consistent with a decrease in gain at the masker place during the course of the masker. It will be assumed that the decrease in gain shown at the signal frequency in Fig. 5 would be occurring at the masker frequency when $\Delta f = 0.4$. In Fig. 8, the two input–output functions from Fig. 5 are replotted. Also shown are data points calculated from the average of the threshold masker levels shown in Fig. 7, when the signal level is 40 dB SPL, assuming the ERB is 517 Hz as above. As in Fig. 7, the open symbol is for the masker in the short-delay condition, and the filled symbol is for the masker in the long-delay condition. Note that the output level is nearly constant in these two conditions. The results are consistent with the idea that the masker must be increased (shown by the horizontal arrow) to produce the same amount of suppression in the long-delay condition as in the short-delay condition. Although the masker output level would only have been decreased by about 9 dB between the two conditions (shown by the downward arrow), the masker level must be increased by 19 dB due to the slope of the input–output function.

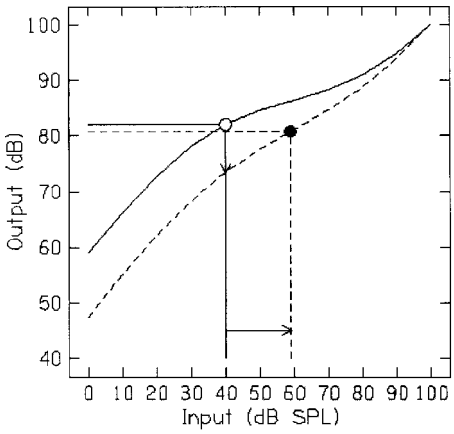


FIG. 8. Schematic illustrating that when average threshold masker levels from Fig. 7 for the 40-dB SPL signal are plotted on the average gain functions from Fig. 5, the masker output levels are nearly constant in the short-delay and long-delay conditions.

3. Other notch widths

The results were analyzed at $\Delta f = 0.0$ and $\Delta f = 0.4$ because of relative computational simplicity. For $\Delta f = 0.0$, there are psychophysical estimates of the signal input–output function that can be used in fitting the data, and the results are consistent with changes in $G_{\max}$ at the signal place. For $\Delta f = 0.4$, it can be assumed that the masker excitation at the signal place is linear, and the results are consistent with a change in $G_{\max}$ at the masker place. At notch widths in between, the results likely reflect changes in $G_{\max}$ at both the signal and the masker frequency. This would require a more comprehensive model to analyze.

IV. GENERAL DISCUSSION

The analyses show that the results would be consistent with a frequency-specific decrease in gain at the masker frequency during the course of masker stimulation. This section will examine the relationship between this decrease in gain and frequency selectivity, previous proposals of this hypothesis, and the physiological mechanism that could underlie this effect.

A. Frequency selectivity

A persistent question in research on the temporal effect has been how to reconcile evidence that gain may decrease during the course of the masker with evidence that frequency selectivity increases with signal delay. As shown in Fig. 3, the change in frequency selectivity is most dramatic at the

lowest signal level. Analysis in terms of input–output functions demonstrates how this could occur. In the left panel of Fig. 5, when $\Delta f = 0.0$, the decrease in G_{max} between the short-delay input–output function (solid line) and the long-delay one (dashed line) causes a decrease in O_{S+M} as well as O_M, and the masker input level does not need to be increased between the short-delay and long-delay conditions. In Fig. 8, when $\Delta f = 0.4$, only the masker output level is decreased in the long-delay condition, and the masker input level must be increased between the short-delay and long-delay conditions. This means that the difference between the masker level at threshold for these two notch widths will increase in the long-delay condition, an indication of sharper frequency selectivity. Heinz *et al.* (2002) show the effects of eliminating suppression from their model while leaving signal gain unchanged, and demonstrate the increase in frequency selectivity seen in the present data.

Frequency selectivity is consistently sharper in the long-delay than the short-delay condition even at higher levels because the temporal effect is larger when $\Delta f = 0.1$ than when $\Delta f = 0.0$ (see Figs. 2 and 3). Although $\Delta f = 0.1$ was not examined in terms of input–output functions, the effects of suppression tend to peak for suppressors between $\Delta f = 0.1$ and 0.2 above the signal frequency (e.g., Sachs and Kiang, 1968; Delgutte, 1990). For this notch width, the temporal effect is most likely a combination of a decrease in gain at the signal frequency and a decrease in suppression.

Some previous studies have interpreted the increase in frequency selectivity with signal delay as evidence for an attentional explanation for the temporal effect. The argument has been that listeners may monitor multiple filters initially, and then narrow their attention to one filter (e.g., Wright and Dai, 1994; Bacon *et al.*, 2002). This would not explain the fact that the amount of change in the filter bandwidth depends on frequency and level, as shown in Table I.

B. Frequency-specific changes in gain

The hypothesis that gain at the suppressor frequency may decrease independently from gain at the signal frequency, as schematized in Fig. 8, is very similar to a hypothesis called ''frequency-specific adaptation'' proposed by Viemeister and Bacon (1982). This hypothesis was proposed to explain masker enhancement, an effect which is probably closely related to the temporal effect. Basically, in masker enhancement, the short-delay and long-delay conditions used in this paper are modified so that the tone is used as a forward masker. The masking produced by the tone is greater in the long-delay condition than in the short-delay condition, suggesting that the gain at the frequency of the tone is actually increased in the long-delay condition. Viemeister and Bacon proposed that the frequency components surrounding the tone suppressed it in the short-delay condition, and in the long-delay condition the suppression adapted.

In the present study, this hypothesis is modified. It is proposed that the gain at the suppressor place is decreased in the long-delay condition, and that this causes a decrease in suppression. As shown in Fig. 8, if the masker level were held constant in the long-delay condition, as in the masker enhancement experiment, the output level of the masker would decrease, and gain at the signal frequency would in fact increase. Figure 8 also demonstrates how a modest change in suppression could produce a large change in threshold, which caused Wright (1995a) to question this hypothesis. As shown in Fig. 8, the change in masker threshold between the short-delay and long-delay conditions depends not only on the amount of change in the masker output level, but also on the slope of the input–output function at the masker frequency. If this slope is shallow, a fairly small change in the masker output level may require a much larger change in the masker input level. This has also been pointed out by Bacon (1996) and Oxenham and Plack (1998).

This hypothesis that suppression may decrease with suppressor duration is consistent with forward masking data showing that the effectiveness of a suppressor decreases as its onset precedes masker onset (Viemeister and Bacon, 1982; Thibodeau *et al.*, 1991; Champlin and Wright, 1993). Physiological data also show that suppression decreases over the first few tens of milliseconds of the suppressor tone (Kiang *et al.*, 1965; Arthur *et al.*, 1971).

C. A possible physiological explanation for the temporal effect

The temporal effect could be mediated by the medial olivocochlear bundle (MOCB) of the efferent system. The MOCB feeds back to the outer hair cells from the level of the superior olivary complex (Warr and Guinan, 1979; Warr, 1980), and appears to decrease amplification during the course of acoustic stimulation. The MOCB feeds back in a frequency-specific manner, so that a fiber with a given CF feeds back to OHCs just basal to the place in the cochlea most responsive to that frequency (Warr, 1992). Efferent activation has mostly been studied using contralateral acoustic stimulation, to separate it from other peripheral effects. Although there are few behavioral studies, quiet threshold for tones has been shown to be elevated by contralateral noise in humans (Strickland and Viemeister, 1995) and macaques (Smith *et al.*, 2000). In the macaques this threshold shift disappeared after section of the MOCB. Contralateral noise has also been shown to decrease the temporal effect in humans (Turner and Doherty, 1997; Bacon and Liu, 2000). May and McQuone (1995) demonstrated an ipsilateral role for MOCB feedback in cats, showing that intensity discrimination for high-frequency tones in noise worsened after MOCB section. Recent studies have shown a change in the level of distortion product otoacoustic emissions (DPOAE) over time in animals that disappears with section of the MOCB (Kujawa and Liberman, 2001).

If feedback from the MOCB decreases the excitation to a sustained stimulus over time, this would result in a decrease in neural firing. This sounds much like neural adaptation, which has long been advanced and rejected as a potential explanation for the temporal effect (Champlin and McFadden, 1989; Bacon, 1990; McFadden and Champlin, 1990). The present data suggest a modification to that hypothesis. It is known that neural adaptation has more than one time constant (Westerman and Smith, 1984), and it could be that only one of these time constants is related to a decrease in gain caused by MOCB feedback, as was noted by

Liberman *et al.* (1996). Also, VIIIth-nerve responses are obtained under anesthesia. Studies show that the change in DPOAE level due to MOCB feedback is greatest in awake animals, and decreases with certain types of anesthesia (Boyev *et al.*, 2002). Therefore, the VIIIth-nerve data may not reflect the efferent feedback that would be seen in awake animals.

D. Implications

Taken together, these results have important implications for hearing. The estimates of frequency selectivity provide a picture of auditory processing under different conditions. When there is an ongoing broadband sound, as when $\Delta f = 0.0$, the system appears to decrease the gain so that over a certain range of levels a tonal signal is more detectable. When a sustained sound is in one frequency region, the system appears to decrease the gain so that sounds in neighboring frequency regions are more audible. This suggests that the auditory system, like other sensory systems, is designed to detect changes in stimuli. The analyses in this paper suggest that this may occur as early as the periphery.

ACKNOWLEDGMENTS

The comments of the editor and two anonymous reviewers were extremely helpful in framing the ideas discussed in this paper. Thanks to Richard Baker for providing help with POLYFIT. This research was partially supported by a grant from NIH (No. R03 DC03510).

[1]It has been argued above that the filter fits to simultaneous notched-noise data do not give an accurate estimate of the excitation in the cochlea. It was decided that the ERB from the middle signal level for the long-duration condition would be used for all signal levels and conditions.

[2]The estimates from the data of Baker *et al.* (1998) were used because they showed a fairly regular increase in G_{max} with frequency, while the estimates from the Glasberg and Moore (2000) data in the paper showed a higher G_{max} for 1 kHz than for 4 kHz. In addition, it should be noted that because Eqs. (2), (3), and (4) are modified from those used by Glasberg and Moore, the G_{max} values in this study may not be directly compared to the ones found in their study.

[3]In this paper, the fits are done as if G_{max} were constant in the long-delay condition. In reality, if G_{max} depends on the level of the masker preceding the signal, it would decrease with threshold masker level, as has been shown by Strickland (2001a).

[4]This estimate was obtained by calculating the threshold output level for the short-duration signal for each listener, using individual quiet thresholds and input–output functions. Assuming the criterion O_{S+M}-O_M is approximately 9 dB at the 4-kHz signal, as above, the masker output level must be at least equal to the threshold output value minus 9 dB. This value (shown by the dotted line) was converted to the masker spectrum level using the same estimate of the ERB at 4 kHz as above, 517 Hz.

Arthur, R. M., Pfeiffer, R. R., and Suga, N. (**1971**). ''Properties of 'two-tone inhibition' in primary auditory neurones,'' J. Physiol. (London) **212**, 593–609.

Bacon, S. P. (**1990**). ''Effect of masker level on overshoot,'' J. Acoust. Soc. Am. **88**, 698–702.

Bacon, S. P. (**1996**). Comments on ''Manipulations of the duration and relative onsets of two-tone forward maskers'' [J. Acoust. Soc. Am. **94**, 1269–1274 (1993)], J. Acoust. Soc. Am. **99**, 3246–3248.

Bacon, S. P., Hedrick, M. S., and Grantham, D. W. (**1988**). ''Temporal effects in simultaneous pure-tone masking in subjects with high-frequency sensorineural hearing loss,'' Audiology **27**, 313–323.

Bacon, S. P., and Liu, L. (**2000**). ''Effects of ipsilateral and contralateral precursors on overshoot,'' J. Acoust. Soc. Am. **108**, 1811–1818.

Bacon, S. P., and Moore, B. C. J. (**1986**). ''Temporal effects in simultaneous pure-tone masking: Effects of signal frequency, masker/signal frequency ratio, and masker level,'' Hear. Res. **23**, 257–266.

Bacon, S. P., Repovsch-Duffey, J. L., and Liu, L. (**2002**). ''Effects of signal delay on auditory filter shapes derived from psychophysical tuning curves and notched-noise data obtained in simultaneous masking,'' J. Acoust. Soc. Am. **112**, 227–237.

Bacon, S. P., and Takahashi, G. A. (**1992**). ''Overshoot in normal-hearing and hearing-impaired subjects,'' J. Acoust. Soc. Am. **91**, 2865–2871.

Bacon, S. P., and Viemeister, N. F. (**1985**). ''Simultaneous masking by gated and continuous sinusoidal maskers,'' J. Acoust. Soc. Am. **78**, 1220–1230.

Baker, R. J., Rosen, S., and Darling, A. M. (**1998**). ''An efficient characterisation of human auditory filtering across level and frequency that is also physiologically reasonable,'' in *Psychophysical and Physiological Advances in Hearing*, edited by A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Whurr, London).

Boyev, K. P., Liberman, M. C., and Brown, M. C. (**2002**). ''Effects of anesthesia on efferent-mediated adaptation of the DPOAE,'' J. Assoc. Res. Otolaryngol. **3**, 362–373.

Carlyon, R. P. (**1989**). ''Changes in the masked thresholds of brief tones produced by prior bursts of noise,'' Hear. Res. **41**, 223–235.

Champlin, C. A., and McFadden, D. (**1989**). ''Reductions in overshoot following intense sound exposures,'' J. Acoust. Soc. Am. **85**, 2005–2011.

Champlin, C. A., and Wright, B. A. (**1993**). ''Manipulations of the duration and relative onsets of two-tone forward maskers,'' J. Acoust. Soc. Am. **94**, 1269–1274.

Cooper, N. P. (**1996**). ''Two-tone suppression in cochlear mechanics,'' J. Acoust. Soc. Am. **99**, 3087–3098.

Cooper, N. P., and Rhode, W. S. (**1995**). ''Nonlinear mechanics at the apex of the guinea-pig cochlea,'' Hear. Res. **82**, 225–243.

Cooper, N. P., and Yates, G. K. (**1994**). ''Nonlinear input–output functions derived from the response of guinea-pig cochlear nerve fibers: Variations with characteristic frequency,'' Hear. Res. **78**, 221–234.

Dallos, P., Harris, D. M., Relkin, E., and Cheatham, M. A. (**1980**). ''Two-tone suppression and intermodulation distortions in the cochlea: Effect of outer hair cell lesion,'' in *Psychophysical, Physiological and Behavioural Studies in Hearing*, edited by G. Van den Brink and F. A. Bilsen (Delft University Press, The Netherlands), pp. 242–249.

Delgutte, B. (**1990**). ''Two-tone rate suppression in auditory-nerve fibers: Dependence on suppressor frequency and level,'' Hear. Res. **49**, 225–246.

Glasberg, B. R., and Moore, B. C. J. (**2000**). ''Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise,'' J. Acoust. Soc. Am. **108**, 2318–2328.

Hant, J. J., Strope, B. P., and Alwan, A. A. (**1997**). ''A psychoacoustic model for the noise masking of plosive bursts,'' J. Acoust. Soc. Am. **101**, 2789–2802.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2002**). ''Quantifying the implications of nonlinear cochlear tuning for auditory-filter estimates,'' J. Acoust. Soc. Am. **111**, 996–1011.

Hicks, M. L., and Bacon, S. P. (**1992**). ''Factors influencing temporal effects with notched-noise maskers,'' Hear. Res. **64**, 123–132.

Hicks, M. L., and Bacon, S. P. (**1999a**). ''Effects of aspirin on the psychophysical measures of frequency selectivity, two-tone suppression, and growth of masking,'' J. Acoust. Soc. Am. **106**, 1436–1451.

Hicks, M. L., and Bacon, S. P. (**1999b**). ''Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing,'' J. Acoust. Soc. Am. **105**, 326–338.

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,'' *Monograph No. 35* (MIT, Cambridge, MA).

Kimberley, B. P., Nelson, D. A., and Bacon, S. P. (**1989**). ''Temporal overshoot in simultaneous-masked psychophysical tuning curves from normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **85**(4), 1660–1665.

Kujawa, S., and Liberman, M. C. (**2001**). ''Effects of olivocochlear feedback on distortion product otoacoustic emissions in guinea pig,'' J. Assoc. Res. Otolaryngol. **2**, 268–278.

Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.

Liberman, M. C., Puria, S., and Guinan, Jr., J. J. (**1996**). ''The ipsilaterally evoked olivocochlear reflex causes rapid adaptation of the 2 $f1-f2$ distortion product otoacoustic emission,'' J. Acoust. Soc. Am. **99**, 3572–3584.

May, B. J., and McQuone, S. J. (**1995**). ''Effects of bilateral olivocochlear

lesions on pure-tone intensity discrimination in cats,'' Aud. Neurosci. **1**, 385–400.
McFadden, D., and Champlin, C. A. (**1990**). ''Reductions in overshoot during aspirin use,'' J. Acoust. Soc. Am. **87**, 2634–2642.
Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1996**). ''Detection of decrements and increments in sinusoids at high overall levels,'' J. Acoust. Soc. Am. **99**, 3669–3677.
Moore, B. C. J., Poon, P. W. F., Bacon, S. P., and Glasberg, B. R. (**1987**). ''The temporal course of masking and the auditory filter shape,'' J. Acoust. Soc. Am. **81**(6), 1873–1880.
Nelson, D. A., Schroder, A. C., and Wojtczak, M. (**2001**). ''A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners,'' J. Acoust. Soc. Am. **110**, 2045–2064.
Nuttall, A. L., and Dolan, D. F. (**1993**). ''Two-tone suppression of inner hair cell and basilar-membrane responses in the guinea pig,'' J. Acoust. Soc. Am. **93**, 390–400.
Oxenham, A. J., and Plack, C. J. (**1997**). ''A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,'' J. Acoust. Soc. Am. **101**, 3666–3675.
Oxenham, A. J., and Plack, C. J. (**1998**). ''Suppression and the upward spread of masking,'' J. Acoust. Soc. Am. **104**, 3500–3510.
Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). ''The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold,'' J. Acoust. Soc. Am. **72**, 1788–1803.
Rhode, W. S., and Cooper, N. C. (**1996**). ''Nonlinear mechanics in the apical turn of the chinchilla cochlea *in vivo*,'' Aud. Neurosci. **3**, 101–121.
Robertson, D., and Johnstone, B. M. (**1981**). ''Primary auditory neurons: Nonlinear responses altered without changes in sharp tuning,'' J. Acoust. Soc. Am. **69**, 1096–1098.
Rosen, S., Baker, R. J., and Darling, A. (**1998**). ''Auditory filter nonlinearity at 2 kHz in normal hearing listeners,'' J. Acoust. Soc. Am. **103**, 2539–2550.
Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). ''Basilar-membrane responses to tones at the base of the chinchilla cochlea,'' J. Acoust. Soc. Am. **101**, 2151–2163.
Sachs, M. B., and Kiang, N. Y-S. (**1968**). ''Two-tone inhibition in auditory-nerve fibers,'' J. Acoust. Soc. Am. **43**, 1120–1128.
Scholl, H. (**1962**). ''Das dynamische Verhalten des Gehors bei der Unterteilung des Schallspektrums in Frequenzgruppen,'' Acustica **12**, 101–107.
Smith, D. W., Turner, D. A., and Hensen, M. M. (**2000**). ''Psychophysical correlates of contralateral efferent suppression. I. The role of the medial olivocochlear system in 'central masking' in nonhuman primates,'' J. Acoust. Soc. Am. **107**, 933–941.
Strickland, E. A. (**2001a**). ''Growth of masker level with signal level as a function of preceding stimulation,'' J. Acoust. Soc. Am. **109**, 2464(A).
Strickland, E. A. (**2001b**). ''The relationship between frequency selectivity and overshoot,'' J. Acoust. Soc. Am. **109**, 2062–2073.
Strickland, E. A., and Viemeister, N. F. (**1995**). ''An attempt to find psychophysical evidence for efferent action in humans,'' Abstracts of the Eighteenth Midwinter Research Meeting of the Association for Research in Otolaryngology, 173(A).
Thibodeau, L. M., Champlin, C. A., and Stiritz, L. (**1991**). ''Suppressor duration effects in forward masking,'' J. Acoust. Soc. Am. **90**, 2268(A).
Turner, C. W., and Doherty, K. A. (**1997**). ''Temporal masking and the 'active process' in normal and hearing-impaired listeners,'' in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ), pp. 387–396.
Viemeister, N. F., and Bacon, S. P. (**1982**). ''Forward masking by enhanced components in harmonic complexes,'' J. Acoust. Soc. Am. **71**, 1502–1507.
von Klitzing, R., and Kohlrausch, A. (**1994**). ''Effect of masker level on overshoot in running- and frozen-noise maskers,'' J. Acoust. Soc. Am. **95**, 2192–2201.
Warr, W. B. (**1980**). ''Efferent components of the auditory system,'' Ann. Otol. Rhinol. Laryngol. **90** (Suppl. 74), 114–120.
Warr, W. B. (**1992**). ''Organization of olivocochlear efferent systems in mammals,'' in *The Mammalian Auditory Pathway: Neuroanatomy*, edited by D. B. Webster, A. N. Popper, and R. R. Fay (Springer, New York), pp. 410–448.
Warr, W. B., and Guinan, Jr., J. J. (**1979**). ''Efferent innervation of the organ of Corti: Two separate systems,'' Brain Res. **173**, 152–155.
Westerman, L. A., and Smith, R. L. (**1984**). ''Rapid and short-term adaptation in auditory-nerve responses,'' Hear. Res. **15**, 249–260.
Wojtczak, M., Schroder, A. C., Kong, Y.-Y., and Nelson, D. A. (**2001**). ''The effect of basilar-membrane nonlinearity on the shapes of masking period patterns in normal and impaired hearing,'' J. Acoust. Soc. Am. **109**, 1571–1586.
Wright, B. A. (**1995a**). ''Across-channel contributions to auditory enhancement,'' in *Advances in Hearing Research: Proceedings of the 10th International Symposium on Hearing*, edited by G. A. Manley, G. M. Klump, C. Koppl, H. Fastl, and H. Oeckinghaus (World Scientific, Singapore), pp. 576–588.
Wright, B. A. (**1995b**). ''Detectability of simultaneously masked signals as a function of signal bandwidth for different signal delays,'' J. Acoust. Soc. Am. **98**, 2493–2503.
Wright, B. A., and Dai, H. (**1994**). ''Detection of unexpected tones in gated and continuous maskers,'' J. Acoust. Soc. Am. **95**, 939–948.
Zwicker, E. (**1965**). ''Temporal effects in simultaneous masking by white-noise bursts,'' J. Acoust. Soc. Am. **37**, 653–666.

Effect of number of masking talkers and auditory priming on informational masking in speech recognition

Richard L. Freyman,[a] Uma Balakrishnan, and Karen S. Helfer
Department of Communication Disorders, University of Massachusetts, Amherst, 715 N. Pleasant Street, Room 6 Arnold House, Amherst, Massachusetts 01003

(Received 11 July 2003; revised 18 January 2004; accepted 1 February 2004)

Three experiments investigated factors that influence the creation of and release from informational masking in speech recognition. The target stimuli were nonsense sentences spoken by a female talker. In experiment 1 the masker was a mixture of three, four, six, or ten female talkers, all reciting similar nonsense sentences. Listeners' recognition performance was measured with both target and masker presented from a front loudspeaker (F–F) or with a masker presented from two loudspeakers, with the right leading the front by 4 ms (F–RF). In the latter condition the target and masker appear to be from different locations. This aids recognition performance for one- and two-talker maskers, but not for noise. As the number of masking talkers increased to ten, the improvement in the F–RF condition diminished, but did not disappear. The second experiment investigated whether hearing a preview (prime) of the target sentence before it was presented in masking improved recognition for the last key word, which was not included in the prime. Marked improvements occurred only for the F–F condition with two-talker masking, not for continuous noise or F–RF two-talker masking. The third experiment found that the benefit of priming in the F–F condition was maintained if the prime sentence was spoken by a different talker or even if it was printed and read silently. These results suggest that informational masking can be overcome by factors that improve listeners' auditory attention toward the target. [DOI: 10.1121/1.1689343]

PACS numbers: 43.66.Dc, 43.66.Pn, 43.66.Qp, 43.71.Es [GDK] Pages: 2246–2256

I. INTRODUCTION

Listening to speech in the presence of competing speech is a complex perceptual task that has been the subject of considerable study over many years (e.g., Broadbent, 1952; Cherry, 1953; Duquesnoy, 1983; Yost *et al.*, 1996; Hawley *et al.*, 1999; Arbogast *et al.*, 2002). One of the characteristics of competing speech that should make it less effective than a continuous masker is that it fluctuates over time, both in spectral composition and in amplitude. There are brief pauses in the competing speech at phrase and sentence boundaries, closures during stop consonants, and very weak consonants, such as /f/ and /θ/, that all create instances of reduced masking. In addition, the spectrum of the competing speech fluctuates independently from the spectrum of the target speech. For example, a high-frequency /s/ sound in the interfering speech may be present simultaneously with a lower-frequency vowel sound in the target speech. These spectral and amplitude fluctuations provide the listener with brief but numerous glimpses of the target speech under conditions in which the target-to-masker ratio is favorable. Normal-hearing listeners seem to be able to use these glimpses to help them understand the target speech in the presence of the competition (Peters *et al.*, 1998). In general, research demonstrates that, decibel for decibel, speech maskers are less effective than noise maskers (see Bronkhorst, 2000).

In contrast to this general result, there appear to be conditions in which competing speech produces additional masking processes beyond those existing for noise maskers (Carhart *et al.*, 1969; Freyman *et al.*, 1999, 2001; Brungart, 2001; Hall *et al.*, 2002). Under such conditions, listeners have great difficulty perceptually extracting target speech material from a complex mixture of voices. Carhart *et al.* (1969) used the term ''perceptual masking'' to describe this phenomenon. Borrowing from nonspeech experiments (e.g., Kidd *et al.*, 1994; Oh and Lutfi, 1998) with which this type of speech masking shares several characteristics, we, and others, have more recently used the term ''informational masking.'' The definition of informational masking in speech recognition appears to be quite broad, encompassing features of masking, or release from masking, that cannot be explained in terms of traditional energetic masking. Among these features are unusually shallow slopes of growth of speech recognition performance with increasing S–N ratio (Freyman *et al.*, 1999; Brungart, 2001; Arbogast *et al.*, 2002) and a large release from masking due to horizontal separation of target and masker (Freyman *et al.*, 1999, 2001; Arbogast *et al.*, 2002; Noble and Perrett, 2002). Both of these characteristics were observed in the multitone masking experiments conducted by Kidd *et al.* (1998).

Ultimately, the basis of informational masking in speech recognition may be discovered by identifying factors that overcome informational masking, allowing the listener to perceptually extract a target that is apparently already represented at some levels of the auditory nervous system. Our previous studies (Freyman *et al.*, 1999, 2001) have focused on creating perceived differences in location between target

[a]Electronic mail: rlf@comdis.umass.edu

and masker as a cue for the listener. Specifically, with target and masker produced from a common front (0-deg) location as a reference (the F–F condition), the experimental condition is where a second source of masking from 60 deg to the right is added, with the right loudspeaker leading the front loudspeaker by 4 ms (the F–RF condition). Due to the precedence effect, the masker is heard at a location very close to 60 deg to the right, while the target is heard directly in front. The F–RF condition produces absolutely no advantage in speech recognition when the masker is continuous noise, but can produce a substantial advantage when the masker is one or two additional voices. Perceptually, the problem of finding and following the target speech within a mixture of several voices is resolved when the apparent location of the masker is moved off to the side.

Although it is difficult to specify precisely the conditions under which informational masking of speech occurs, confusability of the target and masker appears to be a critical feature. For example, Brungart and Simpson (2002) found that a great deal more informational masking occurs when target and masking talkers are of the same sex rather than of the opposite sex, presumably because male and female voices are not highly confusable. There also are likely to be variations in the amount of informational masking within talkers of the same sex, depending on as yet unspecified variables, e.g., similarity in fundamental frequency, speaking rate, speech accent, types of speech materials, etc. For example, for target speech produced by a female native speaker of American English, Freyman *et al.* (2001) found differences in the amount of masking produced by different sets of two female talkers. A composite of two Dutch talkers speaking accented English produced less masking than a complex of two native English speakers.

Carhart *et al.* (1975) reported that the amount of perceptual masking is strongly related to the number of masking talkers. They found that perceptual masking grew as the number of masking talkers increased to three, then decreased as the number was increased further. Hall *et al.* (2002) also reported a large amount of masking for two-talker maskers both in adult and child listeners. Brungart *et al.* (2001) found that in diotic listening conditions, two and three masking talkers produced considerably more masking than one masking talker at low S–N ratios, presumably due to increases in both energetic and informational masking. Yost *et al.* (1996) found that a total of three talkers created considerably more difficulty with a divided attention task than a total of two talkers. Moreover, spatial cues were particularly effective in helping to resolve a condition with three voices, as compared to conditions with two voices. Similarly, Freyman *et al.* (1999, 2001) reported much greater masking for a two-talker masker than for either of the individual talkers separately. The increase was substantially larger in the F–F (nonspatial) condition than in the F–RF condition, in which there was a spatial cue.

The effect of perceived spatial separation in the F–RF condition, as well as the effect of number of talkers, may be explained by the auditory attentional processes in which the listener must be engaged to solve the task. Difficulties in focusing and maintaining attention are likely to be greatest for colocated target and maskers (e.g., the F–F condition). Even so, low-level cues for auditory grouping may still allow segregation into multiple speech streams. The listener attempts to selectively attend to the target utterance and ignore the masking utterance(s). However, especially when both target and masking talkers are of the same sex, attention must be paid to the masking utterances to determine whether they are part of the target speech stream. A listener might attend to the beginning of a masking talker's utterance, decide after a short period that it is not the target sentence, shift attention to the target after missing several words, and possibly lose focus again as attention is pulled away by the competing speech. With two masking talkers, there is likely to be even greater competition for attention than with one masking talker. However, as the number of masking talkers increases much further, they may well create mutual masking of one another, appear less like individual speech streams, and compete less with the target for attention. In the F–RF condition, with the masking talkers perceived in a different location from the target talker, it should be easier for the listener to attend to the target. This type of auditory spatial attention has been shown to provide advantages in both response time and accuracy for identification of nonspeech frequency patterns presented within an informational masking background (Arbogast and Kidd, 2001).

As the number of talkers increases, the additional masker waveforms fill in temporal and spectral gaps and increase the amount of energetic masking in both the F–F and F–RF conditions. However, in the F–F condition a substantial informational component may exist which may be nonmonotonically related to number of talkers, as discussed above. Thus, there is a prediction that the effect of number of masking talkers on speech recognition will proceed quite differently in spatial versus nonspatial conditions, and that the difference in performance in the two conditions will narrow considerably for large numbers of talkers. Experiment 1 of the current paper evaluates this prediction. This investigation also will reveal the number of talkers that produces maximum informational masking for the current stimuli, which will be useful in the design of other experiments with these stimuli, including experiments 2 and 3 of the current paper.

The view that the listener's problem in the nonspatial task is one of identifying and maintaining attention on the target suggests that performance will be improved by any manipulation that helps distinguish the target so that sustained attention can be directed toward it. While spatial separation is clearly useful, other cues may also be effective in helping listeners maintain focus on the target. In experiment 2, we evaluated the usefulness of one such cue, namely whether listeners' ability to follow the target within the target–masker complex is improved if the target is presented in quiet just before the masking trial. By hearing a preview of what to listen for, subjects may be better able to focus on the target early in the trial and less likely to have attention drawn away by the maskers. To make any improvement quantifiable, the last of three key words in a nonsense sentence target was omitted from the preview (priming) stimulus and only this last word was scored when the sentence was

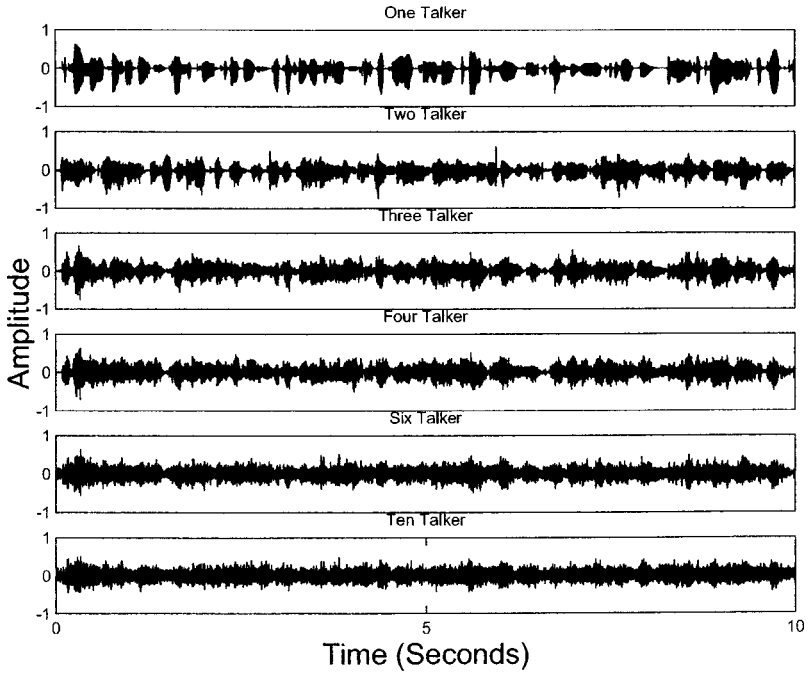


FIG. 1. Ten-second-long segments of single and multitalker maskers.

subsequently presented in masking. Experiment 3 was a follow-up study in which the prime sentence was spoken by a different talker than the target talker or the prime sentence was printed and read by the subject.

II. EXPERIMENT 1: EFFECT OF NUMBER OF TALKERS

A. Method

1. Stimuli

Target stimuli were 320 nonsense sentences spoken by a female talker (Helfer, 1997). These were the same sentences used in experiments described in previous papers (Freyman *et al.*, 1999, 2001). The stimuli had originally been recorded on digital audiotape. The tape recorder's analog output was low-pass filtered at 8.5 kHz and digitally sampled at 20 kHz using a 16-bit A/D converter (TDT AD1). The stimuli were divided into 16 lists of 20 utterances each. Each sentence was semantically incongruous while being syntactically correct, e.g., "The moon could play your love," and contained three key words which were underlined as above for scoring purposes. Percent-correct scores were derived from the number of underlined words correctly identified by the listener.

Four multitalker speech maskers were used: three-talker, four-talker, six-talker, and ten-talker. All maskers were created using the speech of young adult female talkers. Each talker recorded a series of nonsense sentences that was different for each talker and from the 320 target stimuli. The recordings were transferred to a computer (Dell Optiplex GX1p) using a sampling rate of 22.05 kHz. Each talker's recording of discrete nonsense sentences was edited to create an uninterrupted, continuous 35-s-long stream for each talker. The rms outputs of the individual speech streams were equated with one another and then added to build the multitalker maskers as follows: the three-talker masker was created by adding a third talker's speech to the original two talkers (SS and TK) used in Freyman *et al.* (2001), the four-talker masker consisted of the three-talker masker with one more speech stream added to it, and so on. Figure 1 shows 10-s segments of two-, three-, four-, six-, and ten-talker maskers along with the single female talker target for comparison. Note that as the number of talkers in the interference increases, the waveform becomes denser and smoother with a filling in of the peaks and valleys characteristic of the single- and two-talker maskers. Figure 2 displays the long-term one-third-octave spectra of the target and maskers. For ease of viewing, the target was shifted by 20 dB.

2. Apparatus

The experiments were conducted in the same anechoic chamber used for previous experiments (Freyman *et al.* 1999, 2001). It measured $4.9\times4.1\times3.12$ m. The walls, floor, and ceiling are lined with 0.72-m foam wedges. The subject was seated in the center of the room in front of a foam-

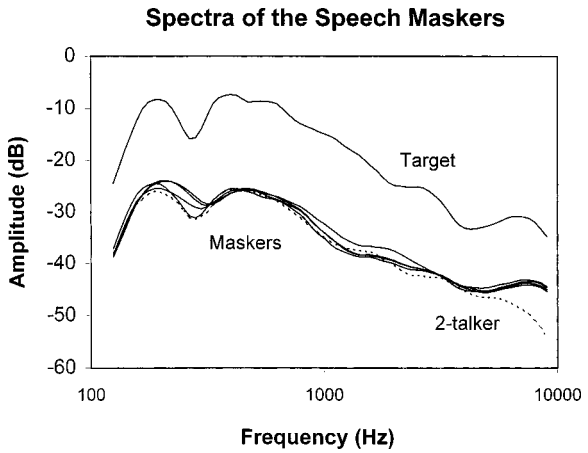


FIG. 2. Long-term one-third-octave spectra of the target and maskers. The target is offset by +20 dB for ease of viewing. The other solid lines represent the spectra of the three-, four-, six-, and ten-talker maskers. The dashed line indicates the spectrum of the two-talker masker.

covered semicircular arc on which two speakers were located. The front loudspeaker was at 0 deg horizontal azimuth and the right speaker was positioned at 60 deg azimuth to the right of the listener. Both were 1.9 m distant from the approximate center of the subjects' heads and at ear height for the typical adult.

The target sentences were delivered via TDT System I instrumentation. The output of the 16-bit D/A converter (TDT DA1) running at 20 kHz was low-pass filtered at 8.5 kHz (TDT), attenuated (TDT PA3), and mixed with the masker before being delivered through a Crown D40 amplifier to a Realistic Minimus 7 loudspeaker. The masker was delivered from the sound board of a Dell Dimension Pentium computer and fed to a delayer (Klark Teknik DN716), one output of which was delayed by 4 ms relative to the other. The delayed output was mixed with the target (TDT SUM3) prior to power amplification and was delivered to the front loudspeaker. The undelayed output was delivered to the right loudspeaker, but was switched off in the conditions in which only the front loudspeaker was to be used. Calibration of presentation level was by means of a 1-in. microphone (B&K 4145) fitted with a random incidence corrector and lowered to the position of the subject's head with the subject absent. A sound-level meter (B&K 2204) located outside the chamber measured the microphone output using the A scale and fast meter response. Small differences measured in the level of the target across the 320 sentences were minimized during the experiment using the TDT PA3 attenuator.

3. Procedures

Two target–masker configurations were used as before: F–F, where the target and masker were presented from the front loudspeaker, and F–RF, where the target was delivered from the front loudspeaker and the masker was delivered from both front and right loudspeakers with a 4-ms time lead to the right. Each of the four maskers was presented at four signal-to-noise (S–N) ratios. The initial data were obtained for the four-talker masker at S–N ratios of −12, −8, −4, and 0 dB. Because of poor performance observed at −12 dB, the S–N ratios for subsequent masker conditions were changed to −8, −4, −2, and 0 dB. The four S–N ratios at two loudspeaker configurations (F–F and F–RF) produced a total of eight conditions per masker.

Specification of target level was based on the median of a large sample of peak needle readings on the sound-level meter. The rms of the maskers (after combining the talkers together) was equated to the rms of a sawtooth wave ($F_0 = 100$ Hz), which was presented daily for calibration. S–N ratios were specified as the difference between the target level of 46 dBA and the measured dBA of the sawtooth wave in the F–F condition. No corrections were made for additional masker energy occurring in the F–RF condition. Across listening blocks of 20 sentences, the desired signal-to-noise ratio was fixed and achieved by varying the level of presentation of the masker while the target level was maintained at 46 dBA.

The subjects were normal-hearing young adults with pure-tone thresholds $\leqslant$20 dB HL in the frequencies 0.5, 1.0, 2.0, 3.0, 4.0, and 6.0 kHz (ANSI, 1996). A different group of eight subjects listened to each masking condition. As in Freyman *et al.* (1999, 2001), a completely within-subjects Latin square design was used for each of the maskers to minimize the potential interaction of subject and sentence list differences. Because there were eight listening conditions for each masker, and 16 lists were available, two consecutive lists were used per condition per listener. Thus, the percentage of key words perceived correctly across subjects for each condition was based on 960 scored items (16 lists $\times$20 sentences$\times$3 key words).

The listener initiated each trial with a button press. The masker was gated on first, with the target sentence following between 0.6 to 1.2 s later, the brief delay in target onset providing a basis for attending to the target. Because the masker was played continuously, its onset during a trial could occur at any point in the continuous speech stream while the target always began with the first word of a nonsense sentence. The target and masker terminated simultaneously. The listener was instructed to repeat the target sentence to the best of his or her ability. While no physical restraints were placed on the listeners, they were advised to maintain a head position facing the front speaker.

Subjects completed the entire listening session in about 1 h, with a break provided halfway through. Prior to listening to the experimental stimuli, subjects listened to five practice sentences to familiarize them with the target speaker's voice. These five sentences were repeated in selected signal-to-noise and speaker conditions to instruct the subject on the task and conditions of the experiment.

B. Results

The basic result of this study is that the improvement in performance in the F–RF condition relative to the F–F condition decreased as the number of masking talkers increased from three to four to six to ten. In the data plotted in panels (b)–(e) of Fig. 3, the narrowing of the difference between the two conditions as the number of talkers is increased is evident. For comparison, the results from the two-talker masker (SS+TK) from experiment 2 of the current paper is displayed in panel (a).[1] There was a considerable narrowing of the F–RF versus F–F difference between the two- and three-talker maskers, and further narrowing as the number of talkers increased to ten.

Signal-to-noise ratios required for a criterion performance of 50% correct were estimated through interpolation of the functions in Fig. 3. Figure 4 [panel (A)] shows the differences in S–N ratio for criterion performance between the F–RF and F–F conditions. In addition to the four maskers studied in the current experiment and the two-talker masker added from experiment 2, the figure also displays single-talker data from Freyman *et al.* (1999, 2001). In Freyman *et al.* (1999), TK was used as a masker, whereas SS was the single-talker masker in Freyman *et al.* (2001). The current figure shows the average of the F–RF versus F–F difference for those two individual talkers. The criterion performance used in these computations for single-talker masking was 60% correct, as subjects never scored as low as 50% at any of the tested S–N ratios. The results show that, among these six conditions, the two-talker masker was associated

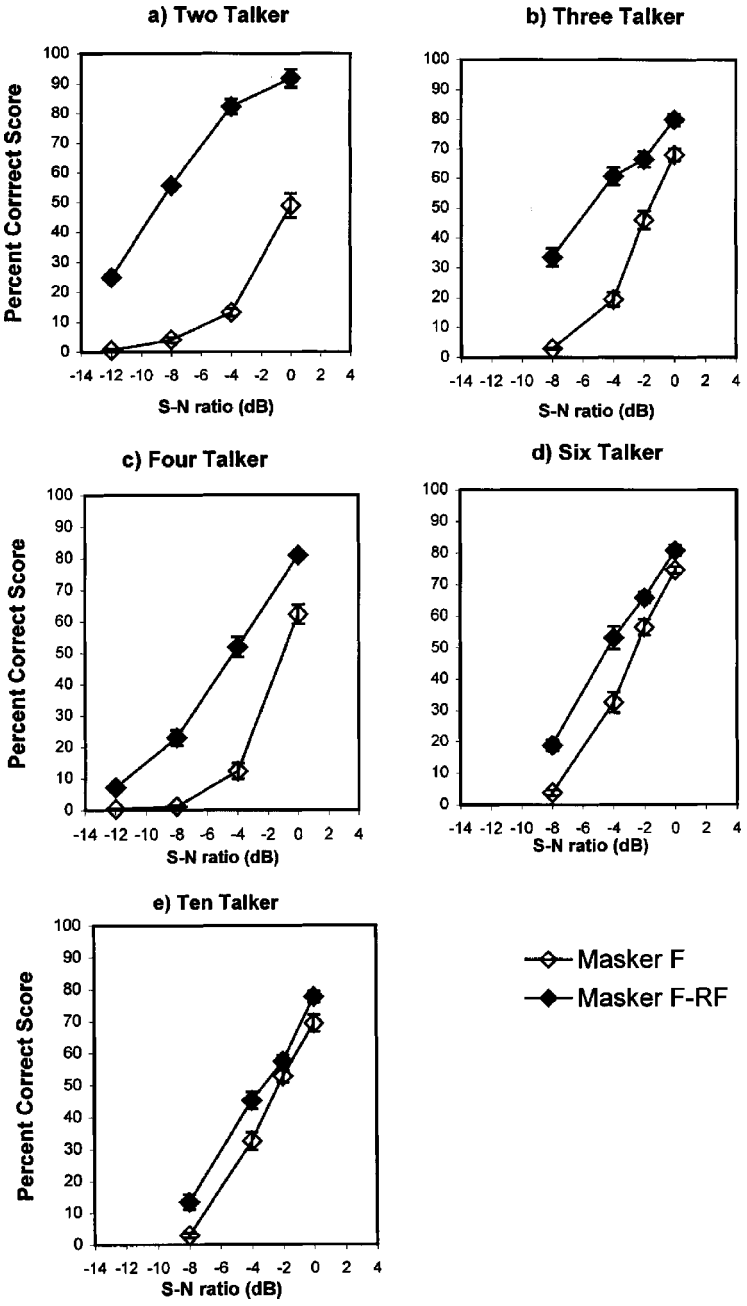


FIG. 3. Group mean-percent correct scores for key words within the target sentences as a function of S–N ratio for the F–F and F–RF conditions. Error bars represent ±1 standard error. Each panel shows the results for a different masker. For comparison, the two-talker masker data obtained in experiment 2 are also included.

with the largest improvement in the F–RF condition relative to the F–F condition. The F–RF benefit was reduced considerably, but did not go to zero, as the number was increased to ten masking talkers. The sources of the narrowing difference are revealed in Fig. 4, panel (B), which shows the actual S–N ratios for criterion performance for the F–RF and F–F conditions individually. The lower line displays the S–N ratio for 50% correct performance for the F–RF condition for the two-, three-, four-, six-, and ten-talker maskers. (The S–N ratio for the single-talker maskers could not be included because specification of S–N ratio was different in the earlier studies and, as described above, performance always was above 50% correct.) The upper line displays the criterion S–N ratios for the F–F condition for these five maskers. This figure shows that the S–N ratio for criterion performance in the F–RF condition, which was presumably largely due to energetic masking, increased as number of masking talkers increased, while the effect was the opposite in the F–F condition. This suggests that the increase in energetic masking over that range (assumed to be the same in F–F and F–RF) was more than offset in the F–F condition by a substantial decrease in informational masking.

III. EXPERIMENT 2: PRIMING BY TARGET TALKER

The first experiment showed that the two-talker masker was most effective in creating informational masking and that perceived differences in spatial location were useful in overcoming this masking, presumably because it facilitated listeners' focused attention on the target. In the current experiment, we explored an alternative means of increasing listeners' ability to identify and focus attention on the target. This experiment investigated the effect of ''priming'' or cuing the listener to the nonsense sentence associated with a

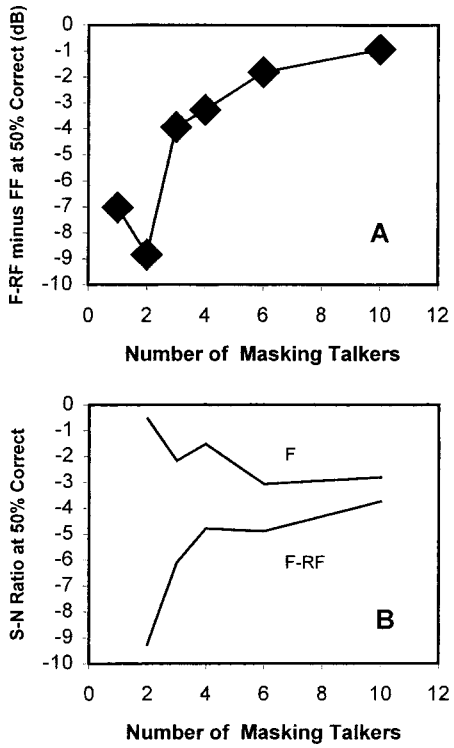


FIG. 4. (A) Difference between the F–F and F–RF conditions for criterion (50%-correct) performance as a function of number of masking talkers. (B) Actual criterion S–N ratios for the F–F and F–RF conditions as a function of number of masking talkers.

target word. In the control condition, the listening task was similar to that described for experiment 1. That is, the listeners heard and repeated back the target nonsense sentences in the presence of masking noise or speech at designated S–N ratios. The comparison, priming condition was identical to the control except that the listener heard the target sentence in quiet just before it was presented in masking. In this preview the last of the three key words in the target sentence was omitted and replaced by noise. Our hypothesis was that hearing the prime would improve listeners' ability to identify and attend to the target utterance when it was presented in a two-talker masker, and therefore improve recognition of the last key word, even though it was not heard during the prime. In continuous noise masking, simple audibility of the target, not attention, is assumed to be the most important factor. Because the last key word could not be predicted from the preceding words in each nonsense sentence, it was hypothesized that the prime would provide no advantage for continuous noise masking.

A. Method

The target sentences were the same set of 320 stimuli used in experiment 1. The priming utterances that preceded these sentences were identical to the target sentences except that the final key word of each utterance was replaced by a noise segment. The noise segment was produced by creating a white-noise token whose duration (700 ms) matched that of the longest third key word segment across all target utterances. The noise was scaled to an rms of approximately 10 dB below the rms of the target speech and appended to the end of each sentence, whose last word had been removed through waveform editing. Figure 5 displays an example target utterance and the corresponding prime utterance.

Two maskers were used: the two-female talker (SSTK) masker used in Freyman *et al.* (2001) and a Gaussian noise whose spectral shape was modeled after filter characteristics described for female speakers of midwestern (Standard) American English (Byrne *et al.*, 1994). Loudspeaker locations and calibration of targets and maskers were as described previously in experiment 1. In the priming condition, an individual trial consisted of the priming utterance (the one with the noise segment at its end) presented in quiet first and followed, after a button press, by the complete target utterance presented against the background masker. Both priming

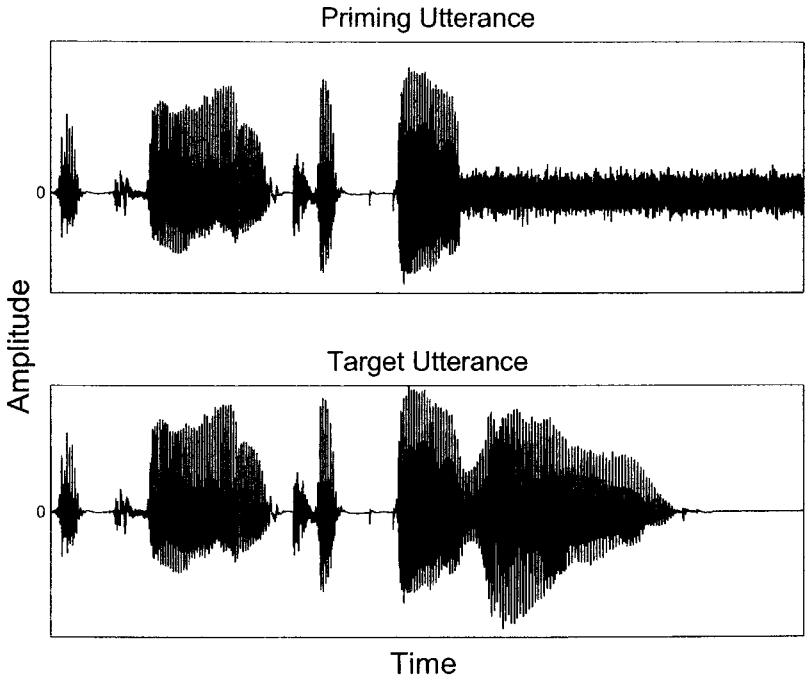


FIG. 5. Time-domain waveforms of the nonsense utterance ''A corn took their wire'' in the priming and target conditions for experiment 2. In the priming condition, the word ''wire'' was replaced by a 700-ms-long segment of white noise.

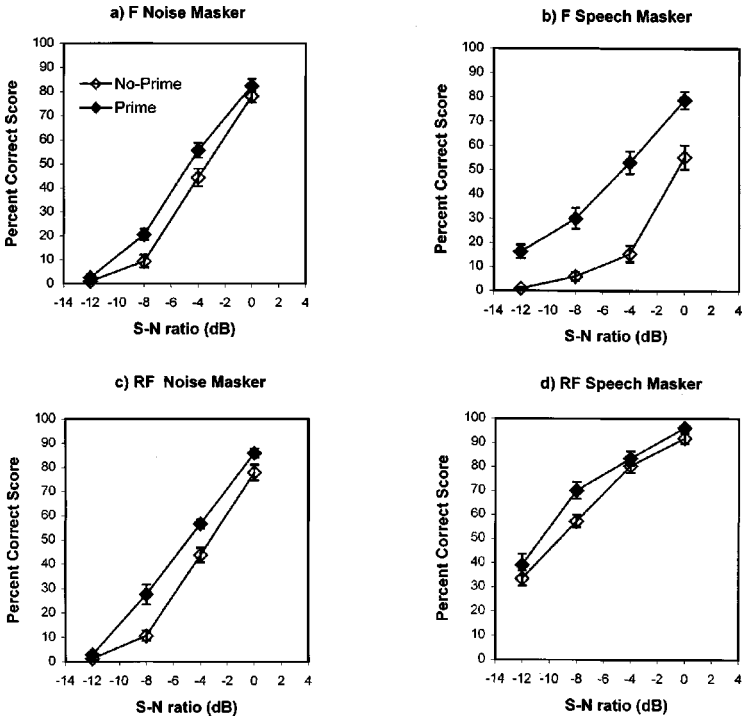


FIG. 6. Effect of priming and spatial separation for a speech-shaped noise masker and two-talker speech masker. Group mean data are shown along with ±1 standard error. Data for the F–F condition were obtained from a different group of subjects than those used for the F–RF condition.

and target utterances were presented at 46 dBA. The listener's task as always was to repeat the entire target sentence, although only the final key word was scored. The no-prime condition was the same except the prime was not presented. Two different groups of 16 subjects participated, one of which listened only to the F–F conditions and the other to the F–RF conditions only. Within the F–F or F–RF configuration, each subject listened to 16 conditions (2 maskers ×2 priming conditions×4 S–N ratios) using a Latin square design in which each condition and subject was assigned all 16 lists, and the condition to list assignment was never repeated across subjects.

B. Results

Figure 6 displays the results for the F–F and F–RF conditions for the noise and speech maskers. Data shown represent the overall percentage-correct score for the last key word in each sentence. Thus, each data point is based on 320 responses (20 items per list×16 subjects/lists). The two functions within each figure, as well as the two figures in the same row, reflect data from one subject group and thus are within-subject comparisons. Comparisons of data across the two rows are between subjects. The results indicate that performance improved with the addition of the priming utterance for every condition. However, the improvement was greatest by far in the F–F two-talker condition where, at the −4-dB S–N ratio, scores improved from 15% correct in the no-prime condition to 53% correct in the prime condition. The improvements displayed in the other three panels at the −4-dB S–N ratio ranged from 3 to 13 percentage points.

As in experiment 1, interpolation of the functions was conducted to derive the S–N ratios at which 50%-correct performance was achieved. Table I displays these derived S–N ratios. Table II displays differences in these S–N ratios for the two maskers to highlight the effect of the priming and F–RF conditions. For the noise masker in both F–F and F–RF speaker configurations and for the two-talker masker in the F–RF speaker configuration, the addition of the prime reduced the criterion S–N ratio by a similar small degree (1.3 to 1.6 dB). The fact that the F–RF speech masker produced priming benefits similar to those obtained with the noise masker suggests that the F–RF two-talker competition, like the noise condition, produced purely energetic masking. That is, it appears that informational masking in the two-talker case was effectively eliminated by spatial separation. It is not clear why performance improved at all in these conditions, as the last key word was never heard in the prime and no semantic context was inherent in the nonsense sentences. It is possible that some phonetic context was provided for the target word by spectral transitions related to coarticulation at the end of the word preceding the (removed) last key word. Another possibility is that the prime decreased the memory load required for the first two words and allowed more resources to be brought to processing of the final word. For the speech masker in the F–F condition, improvement due to the prime was considerably larger, equivalent to an approximately 4-dB reduction in S–N ratio at 50% correct [see also panel (b), Fig. 6]. Here, we assume that informational masking was partially released. Although it gave no direct information about the key word, the priming sentence may have helped the listener to extract the target auditory ''object'' out of the mixture of three talkers. Once the object was extracted and attended to, the last key word was more

TABLE I. S–N ratios (dB) derived by interpolation for 50%-correct response for experiment 2.

Masker	F–F no prime	F–F prime	F–RF no prime	F–RF prime
Noise	−3.33	−4.64	−3.27	−4.90
Two talker	−0.53	−4.54	−9.21	−10.59

TABLE II. Advantage of priming and of spatial separation in dB for the noise and speech maskers in experiment 2. These values were derived from the S–N ratios reported in Table I. The benefit of priming can be seen for the different spatial conditions in the left half of the table. The right half of the table shows the benefit of spatial separation for the no-prime and prime conditions.

Benefit of Priming (dB)			Benefit of F–RF versus F–F (dB)		
	F–F	F–RF	No Prime	Prime	
Noise	1.31	1.63	−0.06	0.26	Noise
Speech	4.01	1.38	8.68	6.1	Speech

understandable because it was connected to that object.

A comparison of the benefit of priming versus spatial separation in Table II suggests that perceived spatial separation was more effective in release from informational masking than was priming. The 4-dB improvement due to priming for the speech masker was not as large as the 8.7-dB improvement obtained in the F–RF condition. Further, in the F–RF condition, the effect of priming was small (1.4 dB), while in the priming condition, the effect of spatial separation was substantial (6.1 dB). This implies a considerable additional release from informational masking in a condition where there was already some release due to priming.

IV. EXPERIMENT 3: COMPARISON OF TYPES OF PRIMING STIMULI

The fact that priming resulted in release from informational masking led to the question of what features of the priming utterance were important for cuing the listener. Because both the priming and target sentences were spoken by the same person, it was possible that one salient cue was the voice and delivery characteristics of the speaker. On the other hand, it could also have been the case that listeners were helped by the priming utterance because they were able to attend to the specific words that had just been presented in the prime. In the next experiment, we varied the priming stimulus to try to distinguish between these possibilities. Because the effect of priming was more robust in the F–F condition, only the F–F condition was used for this experiment.

A. Method

Three priming conditions were used. The first was the same as in the previous experiment; that is, the target speaker's utterance was used as the priming and test utterance (the ''target-talker'' condition). The second condition consisted of the same priming sentences recorded by a young adult male talker (''male talker''). The processing of this prime was identical to that of the target-talker prime, except that the noise segment substituted for the last key word was slightly longer, at 715 ms, in order to match the longest last key word within his recordings of the 320 sentences. The third priming condition (''reading'') consisted of the priming sentences presented in print form with the last word omitted. At the start of each reading prime block, the subject was provided with a set of 20 utterances typed out on index cards with blank cards following each utterance card. The subject was instructed to read the priming utterance, turn that card over to reveal a blank card, and then press a button to listen to the complete target sentence presented with the masker. For comparison, a fourth, no-prime condition was also included.

Only the two-talker speech masker from the previous experiment was used. Hence, there was a total of 16 conditions (4 priming conditions$\times$1 masker$\times$4 S–N ratios). A new group of 16 young normal-hearing subjects was presented with the conditions in a Latin square design as described previously. Signals and maskers were calibrated and presented as described earlier.

B. Results

Figure 7 displays the mean percent-correct scores for the three priming conditions and the no-prime condition. It is apparent that the availability of all three priming conditions improved performance relative to the no-prime condition by approximately the same amount. The two dashed lines replot the results for the target prime and no-prime F–F only conditions from experiment 2, which were obtained with a different group of listeners. These conditions were identical to the target-talker condition of the present study. As can be seen, the effects of priming are consistent across subject groups. See Table III.

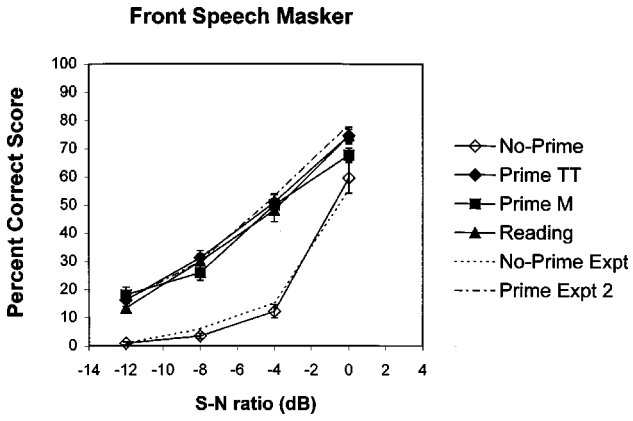


FIG. 7. Comparison of group mean-percent correct scores (with $\pm$1 standard error) for different priming conditions used in experiment 3. The control condition was the ''no-prime'' condition, in which the subjects received no priming utterance (open diamonds). ''Prime TT'' (filled diamonds) refers to the condition where the target talker produced the priming utterance. ''Prime M'' (filled squares) is the condition where the priming utterance was produced by a male talker. ''Reading'' (filled triangles) refers to the priming utterance being presented in print. Dashed lines show the primed and unprimed percent-correct scores obtained in experiment 2.

TABLE III. S–N ratio (dB) derived by interpolation for 50%-correct response for experiment 3. Values obtained for the same conditions in experiment 2 are displayed again to show consistency across subject groups.

	Target talker	Male talker	Reading	No prime
Expt. 3	−4.19	−4.00	−3.72	−0.82
Expt. 2	−4.54	···	···	−0.53

V. DISCUSSION

The first experiment, when considered together with the earlier results presented by Freyman *et al.* (1999, 2001), showed that the release from masking created by the F–RF condition increased when the number of masking talkers was increased from one to two, but decreased as the number was increased further. Performance in the F–F and F–RF conditions followed essentially opposite patterns as the number of talkers increased from two to ten, resulting in a narrowing of the difference between the two conditions. The F–F condition became easier with increasing number of talkers, while the F–RF condition became more difficult. The effect of number of talkers on energetic masking is assumed to be equivalent in the two conditions. Therefore, the different pattern seen in the F–F condition presumably reflects the rise, then fall, of informational masking as the number of masking talkers is increased (see Fig. 4). The overall shape of the function is reminiscent of the nonmonotonic effect of number of masking components on detection of a 1000-Hz signal within a multicomponent informational masker (Oh and Lutfi, 1998). The specific peak observed here for the two-talker masking may be unique to the current stimuli and conditions. However, as noted in the Introduction, the literature includes other examples in which two masking talkers caused significant difficulty for the listener (e.g., Hall *et al.*, 2002; Yost *et al.*, 1996; Carhart *et al.*, 1975). One difference is that Carhart *et al.* (1975) found that three masking talkers produced more perceptual masking than two, whereas with the current stimuli three masking talkers produced less informational masking. These differences can presumably be attributed to the specifics of the stimuli and task.

The second experiment showed that hearing a preview of the target before it was presented in a two-talker masker improved speech recognition performance for a tested key word that was not included in the preview. At some S–N ratios, the subjective impression is that the target message stands out much more clearly in the speech background in trials preceded by the priming utterance. Once the target utterance was perceptually pulled out of the mixture of voices, the listener was often able to follow the message well enough to correctly perceive the unprimed last key word. This idea of ''latching on'' to a talker's voice is consistent with Brungart's (2001) finding that, once listeners decide which of two voices is the target (correctly or incorrectly), they persist in responding with that talker's words. The third experiment demonstrated that the priming utterance need not be the exact target waveform. In fact, it is sufficient for the prime to be spoken by an entirely different talker, or even read by the subject from a printed page. The prime gives the subject information about what to listen for, and this knowledge makes it much easier to attend to the target words and ignore the jumbled utterances of the other two talkers.

These results are consistent with the idea that there were sufficient cues in the target plus two-talker waveform for some level of segregation of the target, even in the nonspatial condition. These cues would include those traditionally thought to be important for auditory grouping, such as temporal asynchrony and differences in fundamental frequency between the target and masking messages (Darwin and Carlyon, 1995), in addition to other differences between target and masking speech (e.g., Cherry, 1953; Darwin and Hukin, 2000; Darwin *et al.*, 2003). Despite the fact that the cues are available, the three mixed utterances are difficult to sort out in the brief observation interval. The voices of the masking talkers compete for attention with the target talker. They are in many of the tested conditions at least as loud as the target, come from the same location as the target, and are perceptually similar to the target along some dimensions in that both target and maskers were produced by adult females. This last issue of similarity between target and masker was studied explicitly by Kidd *et al.* (2002) for the identification of nonspeech auditory patterns. They studied the effectiveness of informational maskers that varied in the extent to which they were likely to perceived as auditory streams that could be confused with the masker. Kidd *et al.* (2002) concluded that it was not possible to distinguish between explanations that relied on similarity between target and masker and those that depended on the allocation of attentional resources to maskers that formed their own perceptual streams. Likewise, in the current experiments, the fact that target and maskers are all produced by female talkers could contribute to the difficulty of the task in more than one way. The speech of the masker female talkers may be easily confused with the target. Additionally, and partially because of the similarity, the masking speech attracts the listener's attention.

The nonmonotonic effect of the number of masking talkers on recognition of the target in the F–F condition (experiment 1) also can be considered in terms of similarity and attention. As the number of masker talkers increases and eventually becomes a general babble, the similarity of the masker and target decreases. The target stands out from this background as long as the S–N ratio is sufficient for audibility, i.e., energetic masking is the limiting factor. The attentional demands imposed by the masker might be expected to follow a nonmonotonic course, similar at least qualitatively to the data. Two masking talkers would be expected to necessitate more attentional resources than one masking talker; however, as the number of talkers is increased much further and the individual utterances are less well recognized, the competition for attention is likely to decrease.

Another factor that may influence auditory attention is the relative loudness of competing utterances. It is reasonable to assume that a listener would be more likely to attend to a louder voice. Evidence for this comes from the single-talker masker data of Brungart *et al.* (2001, Fig. 1, top panel). In their conditions in which the masking has been shown to be almost entirely informational, it might be assumed that loudness difference between target and masker in either direction might become a cue for following the target

message. Under this assumption, 0-dB S–N ratio should be most difficult, with improvements occurring when the masker is either louder or softer than the target. However, their data show that performance is relatively unchanged over a region from −12- to 0-dB S–N ratio, then improves sharply as S–N ratio is increased above 0 dB. This suggests that listeners' attention is drawn to the louder of the two messages, which may supersede any benefit that might have resulted from a simple difference in level.

The current data suggest that the relative loudness of the target voice to the individual masking voices might also be important when there is a small number of masking talkers. For example, in Fig. 3(a) (two-talker), performance in the F–F condition is extremely poor until the 0-dB S–N ratio condition is reached. At that S–N ratio, the target is 3 dB higher in level than either of the masking talkers, whereas at all other S–N ratios, the target is about the same as (−1 dB at the −4-dB S–N ratio condition) or below the level of any individual masking voice. As the number of masking talkers increases and the masker is perceived more as a complex babble than individual voices, the loudness difference between the target and any one voice is not likely to be as important.

The current experiments employed conditions that improved the listener's ability to identify and focus attention on the target talker and not on the masking utterances. In the case of the priming conditions, attention to the target stream is made easier because the subject has heard or seen what sentence to listen for. In the case of the F–RF condition, directional information preserved by the precedence effect allows the already-segregated speech streams to be distinctly localized, and this makes it much easier to attend to, and correctly perceive, the target. There is no suggestion in the data that the precedence effect actually creates the initial segregation. This view is consistent with data showing that in general, basic cues for localization, such as interaural time delay (ITD), are not strong cues for auditory segregation. For example, in the ''double-vowel'' experiments reported by Culling and Summerfield (1995), ITD was not sufficient to segregate vowels when it was the only cue. Rather, interaural differences appear to be important for lateralization of signal components that have been segregated by other means (Hill and Darwin, 1996), and may assist in connecting segregated signals across time (Darwin and Hukin, 1999, 2000). The data from experiment 2 show that the F–RF condition produces greater benefit than the prime condition, and creates additional advantages for recognizing sentences that have already received benefit from priming. Thus, although localization cues are considered to be weak cues for sound-source segregation, localization is extremely useful in the task of selectively attending to one message while ignoring others.

In conclusion, the current results suggest that informational masking is most likely to be observed when one must attend to the speech of one person in the presence of one or two nearby conversations. Conditions that allow the listener to better attend to the target will help overcome this type of masking. Knowing most of what the target talker is going to say ahead of time partially releases informational masking, as evidenced by the enhanced recognition of key words that were omitted from a preview of the target sentences. This enhancement is assumed to be due to an improved ability to identify the target message, requiring fewer attentional resources to be devoted to the maskers. As being exposed to a preview of even a subset of the words to be spoken is unrealistic, future work concerned with finding solutions for overcoming informational masking should consider whether simply knowing the topic provides some benefit.

Informational masking appears to be substantially released by conditions that create a perceived difference in horizontal location between target and interfering speech. This type of release from masking may be unavailable to persons who must listen under conditions in which spatial hearing cannot be well exploited. These would include individuals wearing earmuff hearing protection, people who have bilateral hearing losses but are wearing monaural hearing aids, binaurally fitted hearing aid users who have poor ability to localize sound, and most cochlear implant users, who are generally implanted in one ear. For these situations and individuals, alternative methods will be necessary to achieve target/masker distinctions that facilitate focused and sustained attention on the target message.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute for Deafness and other Communicative Disorders (DC01625). The authors would like to thank Cara Caminiti, Hilary Brown, Joni Skinner, and Wendy Levesque for their assistance in data collection for these experiments.

[1]The priming condition in experiment 2 used a two-talker masker but only the last of the three key words was scored and plotted. For comparisons within experiment 2, the control (no-prime) condition was also scored in the same way. However, for the purpose of comparing the two-talker data from experiment 2 with the other data from experiment 1, all three key words in the no-prime two-talker masker condition were scored for each sentence, exactly as in experiment 1. The primary difference between the collection of the two-talker data plotted in Fig. 1(a) and the data plotted in the other panels was that the data in panel (a) were obtained with two different groups of 16 subjects each (one for F–F and one for F–RF). The data for each of the other panels were obtained within different groups of eight subjects. There is considerable confidence in between-group consistency for identical conditions with these stimuli (see Fig. 7 of the current paper). Further, a within-group comparison for this two-talker masker has already been completed [Fig. 3(B) from Freyman *et al.*, 2001] with similar results, although a difference in the specification of S–N ratio makes a direct comparison with the current data difficult.

ANSI (**1996**). ANSI S3.6-1996, ''Specifications for audiometers'' (American National Standards Institute, New York).

Arbogast, T. L., and Kidd, Jr., G. (**2001**). ''Evidence for spatial tuning in informational masking using the probe-signal method,'' J. Acoust. Soc. Am. **108**, 1803–1810.

Arbogast, T. L., Mason, C. R., and Kidd, Jr., G. (**2002**). ''The effect of spatial separation on informational and energetic masking of speech,'' J. Acoust. Soc. Am. **112**, 2086–2098.

Broadbent, D. E. (**1952**). ''Listening to one of two synchronous messages,'' J. Exp. Psychol. **44**, 51–55.

Bronkhorst, A. W. (**2000**). ''The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions,'' Acust. Acta Acust. **86**, 117–128.

Brungart, D. (**2001**). ''Informational and energetic masking effects in the perception of two simultaneous talkers,'' J. Acoust. Soc. Am. **109**, 1101–1109.

Brungart, D., and Simpson, B. (**2002**). ''The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal,'' J. Acoust. Soc. Am. **112**, 664–676.
Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (**2001**). ''Informational and energetic masking effects in the perception of multiple simultaneous talkers,'' J. Acoust. Soc. Am. **110**, 2527–2538.
Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Nasser Kotby M., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meridith, R., Sirimanna, T., Tavarkiladze, G., Frolenkovgi, G. I., Westerman, S., and Ludvigsen, C. (**1994**). ''An international comparison of long-term average speech spectra,'' J. Acoust. Soc. Am. **96**, 2108–2120.
Carhart, R., Johnson, C., and Goodman, J. (**1975**). ''Perceptual masking of spondees by combinations of talkers,'' J. Acoust. Soc. Am. **58**, S35.
Carhart, R., Tillman, T. W., and Greetis, E. S. (**1969**). ''Perceptual masking in multiple sound backgrounds,'' J. Acoust. Soc. Am. **45**, 694–703.
Cherry, E. C. (**1953**). ''Some experiments on the recognition of speech, with one and two ears,'' J. Acoust. Soc. Am. **25**, 975–979.
Culling, J. F., and Summerfield, Q. (**1995**). ''Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay,'' J. Acoust. Soc. Am. **98**, 785–797.
Darwin, C. J., and Carlyon, R. P. (**1995**). ''Auditory grouping,'' in *The Handbook of Perception and Cognition (Hearing)*, edited by B. C. J. Moore (Academic, London).
Darwin, C. J., and Hukin, R. W. (**1999**). ''Auditory objects of attention: The role of interaural time differences,'' J. Exp. Psychol. **25**, 617–629.
Darwin, C. J., and Hukin, R. W. (**2000**). ''Effectiveness of spatial cues, prosody, and talker characteristics in selective attention,'' J. Acoust. Soc. Am. **107**, 970–977.
Darwin, C. J., Brungart, D. S., and Simpson, B. D. (**2003**). ''Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers,'' J. Acoust. Soc. Am. **114**, 2913–2922.
Duquesnoy, A. J. (**1983**). ''Effect of a single interfering noise and interfering speech upon the binaural sentence intelligibility of aged persons,'' J. Acoust. Soc. Am. **74**, 739–743.
Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). ''The role of perceived spatial separation in the unmasking of speech,'' J. Acoust. Soc. Am. **106**, 3578–3588.
Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (**2001**). ''Spatial release from informational masking in speech recognition,'' J. Acoust. Soc. Am. **109**, 2112–2122.
Hall, J. W., Grose, J. H., Buss, E., and Dev, M. B. (**2002**). ''Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children,'' Ear Hear. **23**, 159–165.
Hawley, M. L., Litovsky, R. Y., and Colburn, H. S. (**1999**). ''Intelligibility and localization of speech signals in a multisource environment,'' J. Acoust. Soc. Am. **105**, 3436–3448.
Helfer, K. S. (**1997**). ''Auditory and auditory-visual perception of clear and conversational speech,'' J. Speech Lang. Hear. Res. **40**, 432–443.
Hill, N. I., and Darwin, C. J. (**1996**). ''Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning,'' J. Acoust. Soc. Am. **100**, 2352–2364.
Kidd, Jr., G., Mason, C. R., Deliwala, P. S., Woods, W. S., and Colburn, H. S. (**1994**). ''Reducing informational masking by sound segregation,'' J. Acoust. Soc. Am. **95**, 3475–3480.
Kidd, Jr., G., Mason, C. R., and Arbogast, T. L. (**2002**). ''Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns,'' J. Acoust. Soc. Am. **111**, 1367–1376.
Kidd, Jr., G., Mason, C. R., Rohtla, T. L., and Deliwala, P. S. (**1998**). ''Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns,'' J. Acoust. Soc. Am. **104**, 422–431.
Noble, W., and Perrett, S. (**2002**). ''Hearing speech against spatially separate competing speech versus competing noise,'' Percept. Psychophys. **64**, 1325–1336.
Oh, E. L., and Lutfi, R. A. (**1998**). ''Nonmonotonicity of informational masking,'' J. Acoust. Soc. Am. **104**, 3489–3499.
Peters, R. W., Moore, B. C. J., and Baer, T. (**1998**). ''Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people,'' J. Acoust. Soc. Am. **92**, 3132–3138.
Yost, W. A., Dye, R. H., and Sheft, S. (**1996**). ''A simulated cocktail party with up to three sound sources,'' Percept. Psychophys. **58**, 1026–1036.

Dominance of missing fundamental versus spectrally cued pitch: Individual differences for complex tones with unresolved harmonics

Remco Renken, J. Esther C. Wiersinga-Post,[a] Sonja Tomaskovic, and Hendrikus Duifhuis

Department of Biomedical Engineering, Neuroimaging Center, Graduate School for Behavioural and Cognitive Neurosciences, University of Groningen, Antonius Deusinglaan 2, 9713 AV, Groningen, The Netherlands

(Received 29 September 2003; revised 3 February 2004; accepted 3 February 2004)

In a two-alternative, forced-choice experiment, subjects had to compare the pitches of two sounds, A and B. Each sound was composed of four successive harmonics of a fundamental frequency between 100 to 250 Hz, added in cosine or Schröder phase. The harmonic frequencies of A were lower than those of B; the missing fundamental frequency of A was higher than that of B. The dominance of the missing fundamental versus the spectrally cued pitch—a pitch percept corresponding to spectral components—was measured as a function of n_A, the lowest harmonic in A. The pitch percept is dominated by the missing fundamental if the harmonics are resolved ($n_A < 7$). If the harmonics become unresolved and are added in Schröder phase, the dominance shifts to a spectrally cued pitch ($7 < n_A < 13$; 75% of the subjects). In the cosine phase condition, many subjects could detect the fundamental pitch well into the unresolved harmonic range ($n_A > 20$). For others, the transition was in the realm of partly resolved harmonics. This shows that the temporal envelope modulation of stimuli with only four unresolved harmonics can give a relatively clear fundamental pitch percept. However, this percept varies considerably among subjects. [DOI: 10.1121/1.1690076]

PACS numbers: 43.66.Hg, 43.66.Nm, 43.66.Lj [NFV] Pages: 2257–2263

I. INTRODUCTION

Most periodic complex sounds have a pitch corresponding to their fundamental frequency, f_0, which remains audible if no energy is present at this fundamental frequency. This well-known phenomenon has been studied extensively during the last century and has been referred to as residue pitch (Schouten, 1940), periodicity pitch, virtual pitch (Terhardt, 1974), low pitch (Smoorenburg, 1970), and pitch of the missing fundamental (for reviews see, e.g., De Boer, 1976; Moore, 1982; Houtsma, 1995).

The audibility of this pitch is generally assumed to be restricted to the frequency range below 5 kHz, and the strength of this pitch depends on the harmonic numbers in the sound (Ritsma, 1962; 1967; Plomp, 1967). Low, resolved harmonics contribute most to this pitch percept. High, unresolved harmonics give a much weaker pitch sensation.

In 1962, Ritsma published experiments with three-tone complexes in which the existence region of the pitch of the missing fundamental was measured. Measurements on three subjects showed that a 200-Hz missing fundamental pitch could be heard for complexes up to harmonic numbers of about 20. Ritsma further showed that the strength of this pitch was positively correlated with the strength of the amplitude modulation of the three-tone complexes. These results were compatible with the then-popular theory that the pitch of the missing fundamental corresponds to the periodicity in the temporal fine structure of the waveform produced by the interaction in a single auditory filter of the unresolved frequency components (Schouten, 1940).

Later, it became clear that the missing fundamental pitch is mainly produced by low, resolved frequency components (Ritsma, 1967; Plomp, 1967). Several studies have been published in which changing the relative phases of the frequency components did not influence the pitch percept, contrary to what Schouten's theory would predict (Patterson, 1973; Wightman, 1973a). Even when the components of a two-component sound were presented dichotically (one harmonic to each ear) a fundamental pitch could be heard, which clearly showed that interaction of several harmonics within one auditory filter is not needed to hear the fundamental pitch (Houtsma and Goldstein, 1972). These findings led to the conclusion that the perception of the fundamental pitch must take place at a central level, where information of separate auditory channels from the two ears is combined and used as input for some kind of pattern recognition process (Goldstein, 1973; Wightman, 1973b; Terhardt, 1974).

However, a few decades later, Houtsma and Smurzynski (1990) rediscovered that a relatively weak fundamental pitch can be heard in the unresolved frequency range and that the strength of this (envelope) pitch sensation diminishes when the amplitude modulation in the sound decreases (Mathes and Miller, 1947; Ritsma, 1962; Houtsma and Smurzynski, 1990). Thus, Schouten's idea that unresolved components can be used for fundamental pitch detection was not completely invalid. This is, e.g., supported by studies that show a pitch doubling when the frequency of the envelope modula-

[a]Author to whom correspondence should be addressed. Electronic mail: j.e.c.post@phys.rug.nl

tion is doubled by adding the unresolved harmonics in sine–cosine phase (Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001). Earlier experiments using sinusoidally amplitude modulated noise showed that a pitch corresponding to the frequency of the amplitude modulation can be heard (Burns and Viemeister, 1976; 1981).

Nowadays, the favored theories on pitch perception assume that both frequency analysis of the resolved components and time-pattern analysis of the unresolved components are involved in pitch perception (De Boer, 1956; Moore, 1982; Srulovicz and Goldstein, 1983; Patterson, 1987; Meddis and Hewitt, 1991a,b).

The pitch of complex sounds with relatively few frequency components is ambiguous. Not only can a pitch be heard corresponding to the fundamental frequency, but also the spectral frequencies, especially the spectral edge frequencies, can provide pitch cues. Current experiments were designed to discriminate between a (missing) fundamental pitch percept and a pitch percept based on spectral cues. The experimental setup resembles that of Smoorenburg (1970) and Houtsma and Fleuren (1991). In these studies, pitch behavior of two-tone stimuli is studied, mainly in the resolved harmonic range (the lowest harmonics are below 12). In the present study, we are especially interested in the behavior of pitch perception in the unresolved and partly unresolved harmonic ranges. Note that we use stimuli with four frequency components. Furthermore, we varied the relative phases of the components, to investigate whether cues based on temporal envelope modulation contribute to the fundamental pitch percept.

The main interest of this study is to see whether there is a change in dominance of the fundamental pitch towards a spectrally cued pitch when the harmonic numbers increase and the components become unresolved or at least partly unresolved. Previous experiments have shown that a weak fundamental pitch, based on temporal envelope cues, can be perceived when *many* unresolved frequency components are present in the stimulus (e.g., Guttman and Pruzansky, 1962; Houtsma and Smurzynski, 1990; Kaernbach and Bering, 2001; Moore *et al.*, 1985). Because of the weakness of this envelope pitch, and because our stimuli contain only four frequency components, we expected to find a dominant spectrally cued pitch in the unresolved harmonic range. However, present experiments show that the temporal envelope pitch can be relatively salient for stimuli with only four frequency components.

In this paper, we will use the following terminology: (1) *The pitch of the missing fundamental* (short: fundamental pitch), to indicate a pitch percept corresponding to the fundamental frequency, independent of the underlying mechanism. (2) *Spectrally cued pitch* (short: spectral pitch), to signify that the pitch percept was not at the fundamental frequency. Rather, it correlated to spectral properties of the sound, like spectral edges or mean frequency of the spectrum. (3) *Envelope pitch*, which refers to a pitch percept at the missing fundamental frequency, based on information of temporal envelope modulation, which is present if the stimulus contains unresolved harmonics.

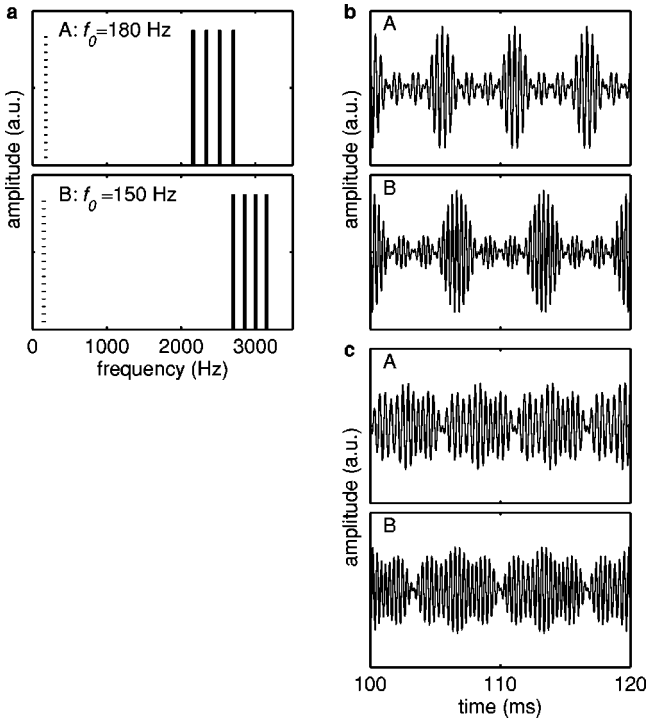


FIG. 1. Example of a stimulus set. (a) Amplitude spectrum of sound A, containing harmonics 12, 13, 14, and 15 of 180 Hz and sound B, containing harmonics 18, 19, 20, and 21 of 150 Hz. (b) Time signals of sounds A and B in which the harmonics are added in cosine phase. (c) Time signals of sounds A and B in which the harmonics are added in Schröder up-chirp phase. The abbreviation a.u. indicates that the corresponding axes are in arbitrary units.

II. METHOD

A two-alternative, forced-choice paradigm was used in which the pitch of two complex sounds, A and B, had to be compared. The sounds A and B were presented diotically and consisted of four harmonics of a missing fundamental. The fundamental frequency of sound A was higher than the fundamental frequency of B, but the harmonics of A were lower in frequency than the harmonics of B. The spectral representation of a stimulus pair is shown in Fig. 1(a).

A and B were randomly presented in the sequence A–B or B–A. The subject's task was to determine whether the second stimulus had the highest or the lowest pitch. This provides information about whether the subject perceived the fundamental pitch (A higher than B) or to the spectrally cued pitch (A lower than B). No feedback was given. Fundamental frequencies of A and B were randomly chosen in the range between 100 and 250 Hz. The difference in fundamental frequency was at least a major second and at most a major third, i.e., their ratio ($f_{0,A}/f_{0,B}$) was between 9/8 and 5/4. Furthermore, the difference in frequency of the highest components ($f_{\max}$) of A and B was at least a major second, i.e., their ratio ($f_{\max,B}/f_{\max,A}$) was at least 9/8, and as close to this value as possible. Thus, all differences were well above discrimination level. The lowest component in sound A was always below the 50th harmonic.

Two stimulus conditions were tested. In condition I, the harmonics of A and B were added in cosine phase. In condition II, the harmonics were added in Schröder "up-chirp"

phase (Schröder, 1970), which minimizes the envelope modulation in the complex sound and reduces the temporal envelope modulation at the level of the basilar membrane (Kohlrausch, 1988). The time signals of the stimuli introduced in Fig. 1(a) are shown in Fig. 1(b) (cosine phase) and Fig. 1(c) (Schröder up-chirp phase). In the remainder of this article we will use the term Schröder phase to refer to the up-chirp Schröder phase condition.

Subjects were seated in a sound-isolated booth (Amplisilence). Stimuli were generated digitally at 44.1 kHz and presented diotically via earphones (Telephonics TDH-49P). Stimuli were presented at 50 dB SPL against a pink-noise background of 40 dB SPL per third octave band. This low stimulus level was used to minimize the influence of combination tones. The duration of the stimuli was 500 ms, with 30-ms cosine ramps. The time between the stimuli was 500 ms. Subjects were forced to give an answer within 1.5 s.

To determine the turning point between a fundamental pitch versus a spectrally cued pitch, the fraction of responses in which stimulus A was indicated as having the highest pitch was measured as a function of the harmonic number in the stimuli using a ''one up''–''one down'' procedure. Three blocks of 60 trials were recorded for both the cosine and Schröder phase. At the start of each block, the lowest harmonic in sound A, n_A, was set at 2. When the subject's answer corresponded to the fundamental pitch (A higher than B), n_A was increased by 1. When the subject's answer corresponded to the spectrally cued pitch (B higher than A) or if no answer was given within 1.5 s, n_A was decreased by 1.

In this experiment, the stimulus with lowest fundamental pitch always contained the highest spectral frequencies. To check whether subjects were using this information, additional control stimuli were presented for which both the fundamental frequency and the spectral frequencies of sound A were higher than those of sound B. On average, 10% of the trials consisted of these control stimuli, which were randomly presented over all trials. The data of these stimuli were analyzed separately. After a trial containing a control stimulus, n_A remained unchanged for the next trial.

A. Data analysis

For three blocks of 60 trials, the fraction of responses in which stimulus A was indicated as having the highest pitch was determined as a function of n_A. There are several methods to obtain the 50% point, $n_{A(P=0.5)}$, i.e., the point along the abscissa at which the fraction indicating stimulus A as having the highest pitch is half (see, e.g., Levitt, 1971). The presented $n_{A(P=0.5)}$ values are derived by fitting a sigmoid (logistic) function to the data using an *N*-weighted maximum likelihood estimation method (with N=number of stimulus presentations; MATLAB toolbox: Wichmann and Hill, 2001a). The high and low asymptotic values could vary between 100%–95% and 0%–5%, respectively. The probability of identifying stimulus A as having the highest pitch—a response corresponding to the fundamental pitch percept—does not drop to chance level with increasing n_A. Rather, subjects' responses will indicate stimulus B as having the highest pitch, which corresponds to a spectrally cued pitch; hence, the low asymptotic value. No slope of the psychometric curve is reported, because the experimental paradigm was not designed to sample the sigmoid completely and therefore does not allow a reliable estimate. The confidence interval of the fitted 50% point was estimated using a Monte Carlo resampling technique relying on 4999 simulated repetitions of the original experiment (Wichmann and Hill, 2001b).

The 50% points, derived by making a sigmoid fit, were compared with the results from a midrun average estimate (see Wetherill, 1963, referenced by Levitt, 1971). The midrun estimate was determined using the last eight runs for each block (data not shown). If an accurate fit was obtained (see below), the 50%-point estimate from the fit was within the range of midrun average values obtained for the three blocks. Therefore, only the 50% point of the sigmoid fit is presented.

In a number of cases, the distribution of data points was rather wide [see Figs. 2(d), (h)]. In these cases, the sigmoid fit becomes rather arbitrary, as can be seen from the estimated confidence intervals (see Fig. 3, subjects 17–23). When no accurate fit was obtained, the fitting procedure showed a tendency to overestimate the $n_{A(P=0.5)}$ value. At the same time, the midrun-average method had a tendency to underestimate this value. Therefore, a lower boundary for the 50% point was estimated by visual inspection of the raw data.

B. Subjects

Twenty-three, out of an initial pool of 30, normal-hearing subjects participated (11 male and 12 female). Their ages ranged from 22 to 39. Most of them (20) had some musical experience. Before starting the experiment, subjects performed two test experiments, one with cosine and one with Schröder phase stimuli, in which there was no limit on the response time. Seven of the initial 30 subjects were excluded from participation. These subjects, all without musical experience, were not able to perform the test experiments, i.e., they performed randomly even when stimuli with low harmonic numbers were presented.

III. RESULTS

Figure 2 shows the results for four differently behaving subjects. The solid and open symbols represent the results for the cosine and Schröder phase condition, respectively. The right column shows n_A as a function of trial number for the first of the three recorded blocks in the cosine (solid symbols) and Schröder (open symbols) phase condition. In the left column, the fraction indicating stimulus A as having the highest pitch—corresponding to a fundamental pitch percept—is plotted as a function of n_A. The area of the symbols is proportional to $N^{1/2}$, with N=number of stimulus presentations.

For all subjects, the probability of indicating stimulus A as having the highest pitch is high at relatively low harmonic numbers ($n_A<7$) and drops with increasing *n*. Due to the experiment protocol, only a few stimuli are presented with either a strong fundamental pitch (stimuli with relatively low

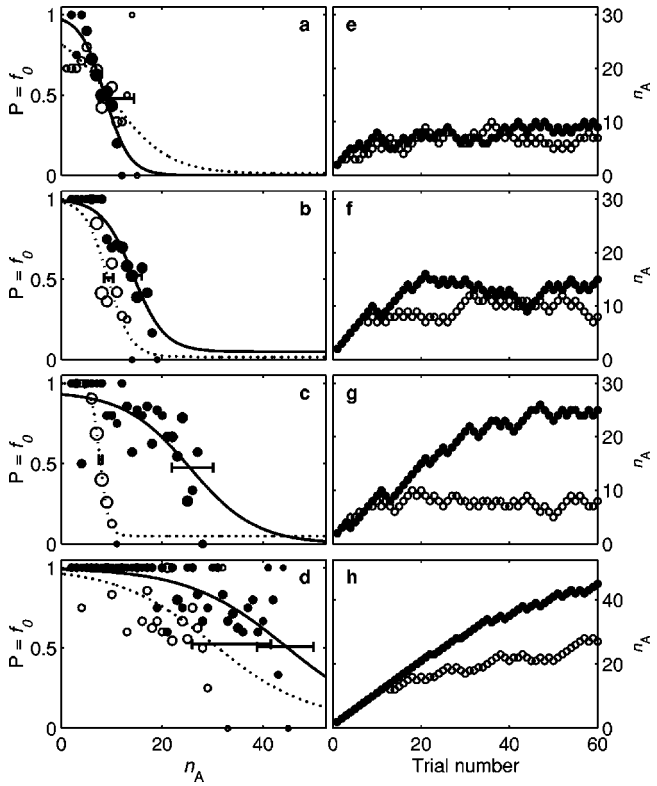


FIG. 2. Left: Psychometric functions of subjects 2 (a), 9 (b), 17 (c), and 20 (d) showing the fraction (P) of responses corresponding to a fundamental pitch percept as a function of n_A, the lowest harmonic number in sound A. Closed symbols: results of cosine phase stimuli; open symbols: results of Schröder phase stimuli. The surface of the symbols is proportional to $N^{1/2}$ (N=number of stimulus presentations). Solid lines: sigmoid fits to the cosine phase data; dotted lines: sigmoid fits to the Schröder phase data. Error bars represent the 95%-confidence intervals of the $n_{A(P=0.5)}$ estimates. Right: n_A as a function of trial number for subjects 2 (e), 9 (f), 17 (g), and 20 (h). Closed symbols: data from the first block in the cosine phase condition; open symbols: data from the first block in the Schröder phase condition.

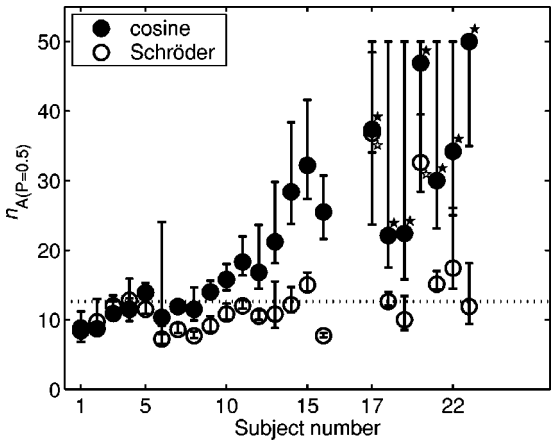


FIG. 3. $n_{A(P=0.5)}$ values with 95%-confidence intervals for all subjects. Closed symbols: cosine phase condition; open symbols: Schröder phase condition. The dotted horizontal line indicates the upper quartile value of $n_{A(P=0.5)}$ for the Schröder phase. For the seven subjects at the right (subjects 17 to 23), the cosine and/or the Schröder phase measurements could not be properly described with a sigmoid function. These data sets are marked with a star.

n) or a strong spectrally cued pitch (stimuli with relatively high n). Most stimuli are located in the region where the pitch is ambiguous.

To determine the harmonic number at which the probability of identifying stimulus A as having the highest pitch is half, $n_{A(P=0.5)}$, sigmoid fits were made. In Fig. 2, the solid lines represent fits to the cosine phase data, and the dotted lines to the Schröder phase data. It should be noted that the paradigm used is adequate in finding $n_{A(P=0.5)}$, but was not designed to give a complete sampling of the sigmoid curve. Hence, the slope of the sigmoid, at $P=0.5$, cannot be estimated reliably from these data sets.

Figure 3 and Table I show the $n_{A(P=0.5)}$ values for all subjects. The error bars in Fig. 3 represent the 95%-confidence intervals. Subjects (1 to 16 and 17 to 23) are ordered according to the absolute difference in the $n_{A(P=0.5)}$ values between cosine and Schröder phase conditions. The $n_{A(P=0.5)}$ values marked with a star (subjects 17 to 23) could not be described adequately using a sigmoid [see, e.g., Fig. 2(d)]. In Table I, we have reported a lower boundary, which was determined by visual inspection of the raw data [see, e.g., Fig. 2(h)].

The first four subjects in Fig. 3 and Table I [see also Fig. 2(a)] obviously show no difference in $n_{A(P=0.5)}$ values between the cosine and Schröder phase conditions ($n_{A(P=0.5)} \approx 10$). The other subjects [subjects 5 to 23, see also Fig. 2 panels (b), (c), and (d)] do show such a difference. For the Schröder phase condition, the $n_{A(P=0.5)}$ values are comparable to those of subjects 1 to 4 and lie within the range of 7 to 17, with an upper quartile value of 13 (see the dotted line in Fig. 3). Subject 20 performed differently and showed an $n_{A(P=0.5)}$ value greater than 30 for the Schröder phase condition [see Figs. 2(d) and (h)]. Overall, the cosine phase condition results in higher $n_{A(P=0.5)}$ values compared to the

TABLE I. $n_{A(P=0.5)}$ values for both cosine and Schröder phase. For some subjects, only a lower boundary for $n_{A(P=0.5)}$ is given (see the text). The fourth column shows the results of the control stimuli.

	$n_{A(P=0.5)}$		control
Subject #	Cosine phase	Schöder phase	% correct
1	8	9	77%
2	9	10	97%
3	11	12	100%
4	12	13	100%
5	14	12	100%
6	10	7	95%
7	12	9	97%
8	12	8	86%
9	14	9	100%
10	16	11	96%
11	18	12	97%
12	17	10	95%
13	21	11	100%
14	28	12	91%
15	32	15	97%
16	26	8	97%
17	>30	>15	82%
18	>17	13	92%
19	>15	10	95%
20	>30	>20	91%
21	>20	15	82%
22	>20	17	97%
23	>30	12	93%

Schröder phase condition. Furthermore, the intersubject variability is higher in the cosine phase condition; values range from 10 to well above 30.

Figure 3 shows that with increasing $n_{A(P=0.5)}$ values, the 95%-confidence intervals increase as well. This is due to the paradigm of the experiment. Because each block started at $n_A=2$ and the number of trials in a block was limited to 60, fewer samples around $n_{A(P=0.5)}$ were presented when the $n_{A(P=0.5)}$ value was high.

The fourth column of Table I shows the results of the control stimuli, in which both the fundamental frequency and the spectral frequencies shift in the same direction. The results are extracted from all six blocks, three blocks of cosine stimuli and three blocks of Schröder stimuli. The percentage correct is clearly above chance level for all subjects, showing that the subjects did not use the information that stimuli with lowest fundamental pitch had the highest spectral components.

IV. DISCUSSION

The setup of the present experiments was such that a pitch percept of the missing fundamental could be distinguished from a pitch percept based on spectral cues. When stimulus B was indicated as having the highest pitch (corresponding to a pitch percept based on spectral cues), it is not clear whether subjects used timbre or other spectral cues for their pitch perception task. Following Plomp and Steeneken, timbre is interpreted as: ''that attribute of auditory sensation in terms of which a listener can judge that two sounds, similarly presented and having the same loudness and pitch are dissimilar'' (Plomp and Steeneken, 1969). Timbre can be ordered on a sharp-to-round scale corresponding to sounds containing relatively many high-frequency components versus relatively many low-frequency components (Plomp, 1976). Thus, timbre becomes sharper if the relative contribution of high frequencies increases. Also, the spectrally cued pitch becomes higher if stimuli contain higher frequency components. Thus, an increase in spectrally cued pitch automatically implies a sharper timbre, which makes the distinction between the two impossible. In this experiment, subjects were asked to base their decisions on pitch cues, and they reported that all their decisions were made based on pitch differences. When subjects responded based on the pitch of the missing fundamental, it can be concluded that they did not focus on timbre. Remember that if the fundamental frequency decreased, spectral components, and thus the sharpness of timbre, increased (and vice versa).

The choice for an adaptive one up, one down approach may lead to hysteresis effects. Subjects may be guided into a ''fundamental listening mode.'' Therefore, our results may show higher transition points than would be obtained if another approach, e.g., starting at high harmonic numbers, was used. However, this does not affect our conclusions that the auditory system is intrinsically capable of identifying the fundamental pitch when stimuli contain unresolved harmonics only, and that the saliency of this fundamental pitch varies greatly among subjects.

A. Pitch perception for resolved harmonics

Current experiments were set up to find out in what harmonic range the pitch of the missing fundamental dominates the pitch percept for stimuli containing only four successive harmonics. For all subjects, the pitch of the missing fundamental dominates the percept at low, resolved harmonics. Comparable experiments with two-tone stimuli (Smoorenburg, 1970; Houtsma and Fleuren, 1991; Hartmann, 1993) showed that many subjects did not hear the fundamental pitch, but listened to the spectrally cued pitch. This discrepancy is probably caused by the difference in the number of frequency components used. Increasing the number of components increases the salience of the fundamental pitch (Houtsma and Goldstein, 1971; Plomp, 1976).

The current results are consistent with literature where it is shown that the relative phases of resolved components do not influence the pitch percept (Patterson, 1973; Wightman, 1973a).

B. Pitch perception for unresolved harmonics

A transition between the fundamental pitch and the spectrally cued pitch percept was expected at the transition from resolved to unresolved harmonics. This expectation was based on the fact that for unresolved harmonics the pitch of the missing fundamental is relatively weak, especially when the stimuli contain only a few harmonics (Houtsma and Smurzynski, 1990). Thus, for stimuli composed of only four unresolved harmonics the spectral frequencies were expected to dominate the pitch percept. Furthermore, the phase relation of low, resolved harmonics does not influence the perception of the fundamental pitch. Therefore, the transition between the missing fundamental and spectrally cued pitch was expected to be independent of the phase relations between the components.

The results show that this expectation is confirmed for some subjects (1 to 4). For these subjects, the $n_{A(P=0.5)}$ values are around 10 and there is no clear difference between the $n_{A(P=0.5)}$ values of the cosine and Schröder phase conditions. However, most subjects do show a difference between the cosine and Schröder phase conditions. For these subjects, the range over which the fundamental pitch is perceived can extend well into the range of unresolved harmonics when the harmonics are added in cosine phase, while the Schröder phase transition stays within the realm of (partly) resolved harmonics. Apparently, the envelope modulation can be the dominant cue for the detection of pitch when harmonics are unresolved and added in cosine phase.[1] That this cue can be used for the detection of the fundamental pitch has been shown in several studies (Guttman and Pruzansky, 1962; Ritsma, 1962; Moore *et al.*, 1985; Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Kaernbach and Bering, 2001). These studies show that the fundamental pitch percept based on cues of envelope modulation is relatively weak. Therefore, many harmonics are presented to elicit this pitch percept in almost all studies. Present experiments show that envelope modulation can elicit a relatively strong pitch (stronger than the spectrally cued pitch) even if the stimuli contain only a few, in this case four, harmonics.

C. Envelope pitch for Schröder phase stimuli

Adding the harmonics in Schröder up-chirp phase strongly reduces the envelope modulation in the stimulus (see Fig. 1) and at the cochlear output (Schröder, 1970; Kohlrausch, 1988). For 75% of the subjects, the $n_{A(P=0.5)}$ value was below the 13th harmonic for the Schröder phase condition, which is around the transition from resolved to unresolved harmonics. This agrees well with the values in literature (Hoekstra, 1979; Houtsma and Smurzynski, 1990). Thus, for most subjects the residual envelope modulation in the Schröder phase condition is too weak to elicit a dominant fundamental (envelope) pitch percept in the unresolved harmonic range. At least one subject (subject 20) could perceive the fundamental pitch of the Schröder phase stimuli well into the unresolved harmonic range. This subject might be very sensitive to envelope modulation cues. It should be noted that the reduction of envelope modulation due to the Schröder phase is probably less effective after cochlear filtering.

D. Envelope pitch for cosine phase stimuli

The result of subject 20 for the Schröder phase already suggests that differences in sensitivity to envelope modulation exist. This is reflected more strongly in the large variability in the $n_{A(P=0.5)}$ values in the cosine phase condition. For several subjects, the transition point for cosine stimuli lies below the 20th harmonic, the range in which the harmonics are still partly resolved (Hoekstra, 1979). For others, this transition point was clearly above the 20th harmonic, and for a few subjects it even reached above the 30th harmonic (see Fig. 3 and Table I).

Apparently, for this last category of subjects, the fundamental pitch was detectable when the frequency of the harmonics exceeded the 4- to 5-kHz range. This ''threshold'' is important in the debate about the mechanism of pitch perception for unresolved harmonics. In this paper, we have used the term ''envelope pitch'' to refer to a fundamental pitch percept based on information of envelope modulation. This term implies that the envelope of the signal is the major contributor to the pitch percept. Physiological measurements show that, at the level of the auditory nerve, phase locking occurs to the temporal envelope (Cariani and Delgutte, 1996; Javel, 1980; Joris and Yin, 1992). However, if the frequencies in the stimulus are low enough, i.e., below the 4- to 5-kHz threshold, phase locking to the temporal fine structure occurs as well (Kiang *et al.*, 1965; Rose *et al.*, 1967). For many subjects, the transition point was below this threshold. For these subjects, it cannot be excluded that temporal fine structure is used as a cue to detect the fundamental pitch (Schouten, 1940; Hall, III *et al.*, 2003). For subjects having a high transition point ($>$30th harmonic), however, phase locking to the temporal fine structure cannot occur (Rose *et al.*, 1967). Therefore, only envelope cues were available to detect the fundamental pitch. Based on these last results, we have opted for the term envelope pitch.

We have reported large differences in sensitivity to envelope modulation, as a cue for fundamental pitch perception, between subjects. Our results are in accordance with those of Burns and Viemeister (Burns and Viemeister, 1981), who used many subjects and found a large intersubject variability in a melody identification task using amplitude-modulated noise. This variability might be caused by differences in individual levels of musical experience. It can explain the controversy that characterizes the discussion on pitch perception in the unresolved harmonic range. Several authors have reported that a fundamental pitch can only be heard for stimuli with relatively low, resolved harmonics ($n<15$; e.g., De Boer, 1956; Houtsma and Goldstein, 1972; Pressnitzer *et al.*, 2001), while others find that a faint fundamental pitch can be heard in the unresolved harmonic range (e.g., Ritsma, 1962; Houtsma and Smurzynski, 1990) and that pitch discrimination, although poorer than in the resolved harmonic range, is still possible (Hoekstra, 1979; Cullen and Long, 1986). In most of these studies only a few subjects participated. If these subjects do not cover the full range of possible behaviors, different outcomes are to be expected. In fact, the current results show that the different outcomes of earlier experiments are probably caused by intersubject variability.

Based on our results, we suggest that the auditory system is intrinsically capable of identifying a pitch at the missing fundamental using envelope modulation. Whether the fundamental pitch is actually perceived depends on the interpretation of the information encoded by the auditory system. Therefore, using a selected, well-trained subject population for pitch-related psychophysical experiments will give information about the capacities of the human auditory system, but not about the mechanism of pitch perception in general.

V. CONCLUSIONS

The current experiments show that sounds composed of a few (four) successive components can elicit a fundamental pitch percept, based on envelope modulation, in the unresolved frequency range ($n>20$). The saliency of the envelope pitch varies greatly among subjects. Some subjects appear to be rather insensitive to the envelope modulation as a cue for fundamental pitch detection, while others can extract this information from stimuli with very high harmonic numbers (>20). This exceeds the harmonic range often reported in literature.

ACKNOWLEDGMENTS

The authors wish to thank Sybrand van der Werf for performing a pilot study and Desiree Houkema for extensive literature research and many fruitful discussions. We thank the editor, N. Viemeister, and the reviewers for their constructive comments and suggestsions.

[1]This effect might have been enhanced if a Schröder ''down-chirp'' phase was used instead of the cosine phase (Kohlrausch, 1988). However, the analysis of these effects within the cochlea is very sensitive to characteristics of the individual ears and of specific cochlea model parameters. This discussion falls beyond the scope of this paper.

Burns, E. M., and Viemeister, N. F. (**1976**). ''Nonspectral pitch [melodic information perception],'' J. Acoust. Soc. Am. **60**, 863–869.

Burns, E. M., and Viemeister, N. F. (**1981**). ''Played-again SAM: Further observations on the pitch of amplitude-modulated noise,'' J. Acoust. Soc. Am. **70**, 1655–1660.

Cariani, P. A., and Delgutte, B. (**1996**). ''Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch,'' J. Neurophysiol. **76**, 1717–1734.
Cullen, J. K., and Long, G. R. (**1986**). ''Rate discrimination of high-pass-filtered pulse trains,'' J. Acoust. Soc. Am. **79**, 114–119.
De Boer, E. (**1956**). ''On the residue in hearing,'' Ph.D. thesis, University of Amsterdam.
De Boer, E. (**1976**). ''On the residue and auditory pitch perception,'' in *Auditory System Clinical and Special Topics*, edited by W. D. Keidel and W. D. Neff (Springer, Berlin), pp. 479–583.
Goldstein, J. L. (**1973**). ''An optimum processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54**, 1496–1516.
Guttman, N., and Pruzansky, S. (**1962**). ''Lower limits of pitch and musical pitch,'' J. Speech Hear. Res. **5**, 207–214.
Hall, III, J. W., Buss, E., and Grose, J. H. (**2003**). ''Modulation rate discrimination for unresolved components: Temporal cues related to fine structure and envelope,'' J. Acoust. Soc. Am. **113**, 986–993.
Hartman, W. M. (**1993**). ''Auditory demonstrations on compact disk for large N,'' J. Acoust. Soc. Am. **93**, 1–16.
Hoekstra, A. (**1979**), ''Frequency discrimination and frequency analysis in hearing,'' Ph.D. thesis, University of Groningen.
Houtsma, A. J. M. (**1995**). ''Pitch perception,'' in *Hearing, Handbook of Perception and Cognition*, 2nd ed., edited by B. C. J. Moore (Academic, San Diego), pp. 267–295.
Houtsma, A. J. M., and Fleuren, J. F. M. (**1991**). ''Analytic and synthetic pitch of two-tone complexes,'' J. Acoust. Soc. Am. **90**, 1674–1676.
Houtsma, A. J. M., and Goldstein, J. L. (**1971**). ''Fundamental tracking and the number of stimulus partials,'' in ''Perception of musical intervals: Evidence for the central origin of the pitch of complex tones,'' MIT RLE Technical Report 484, pp. 24–27.
Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ''The central origin of the pitch of complex tones: Evidence from musical interval recognition,'' J. Acoust. Soc. Am. **51**, 520–529.
Houtsma, A. J. M., and Smurzynski, J. (**1990**). ''Pitch identification and discrimination for complex tones with many harmonics,'' J. Acoust. Soc. Am. **87**, 304–310.
Javel, E. (**1980**). ''Coding of AM tones in the chinchilla auditory nerve: Implications for the pitch of complex tones,'' J. Acoust. Soc. Am. **68**, 133–146.
Joris, P. X., and Yin, T. C. T. (**1992**). ''Responses to amplitude-modulated tones in the auditory nerve of the cat,'' J. Acoust. Soc. Am. **91**, 215–232.
Kaernbach, C., and Bering, C. (**2001**). ''Exploring the temporal mechanism involved in the pitch of unresolved harmonics,'' J. Acoust. Soc. Am. **110**, 1039–1048.
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 Press, Cambridge, MA).
Kohlrausch, A. (**1988**). ''Masking patterns of harmonic complex tone maskers and the role of the inner ear transfer function,'' in *Basic Issues in Hearing, Proceedings of the 8th International Symposium on Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, London), pp. 339–350.
Levitt, H. (**1971**). ''Transformed up–down methods in psychoacoustics,'' J. Acoust. Soc. Am. **49**, 467–477.
Mathes, R. C., and Miller, R. L. (**1947**). ''Phase effects in monaural perception,'' J. Acoust. Soc. Am. **19**, 780–797.
Meddis, R., and Hewitt, M. J. (**1991a**). ''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. (**1991b**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity,'' J. Acoust. Soc. Am. **89**, 2883–2894.
Moore, B. C. J. (**1982**). *An Introduction to the Psychology of Hearing* (Academic, London).
Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). ''Relative dominance of individual partials in determining the pitch of complex tones,'' J. Acoust. Soc. Am. **77**, 1853–1860.
Patterson, R. D. (**1973**). ''The effects of relative phase and the number of components on residue pitch,'' J. Acoust. Soc. Am. **53**, 1565–1572.
Patterson, R. D. (**1987**). ''A pulse ribbon model of monaural phase perception,'' J. Acoust. Soc. Am. **82**, 1560–1586.
Plomp, R. (**1967**). ''Pitch of complex tones,'' J. Acoust. Soc. Am. **41**, 1526–1533.
Plomp, R. (**1976**). *Aspects of Tone Sensation* (Academic, London).
Plomp, R., and Steeneken, H. J. M. (**1969**). ''Effect of phase on the timbre of complex tones,'' J. Acoust. Soc. Am. **46**, 409–421.
Pressnitzer, D., Patterson, R. D., and Krumbholz, K. (**2001**). ''The lower limit of melodic pitch,'' J. Acoust. Soc. Am. **109**, 2074–2084.
Ritsma, R. J. (**1962**). ''Existence region of the tonal residue. I,'' J. Acoust. Soc. Am. **34**, 1224–1229.
Ritsma, R. J. (**1967**). ''Frequencies dominant in the perception of the pitch of complex sounds,'' J. Acoust. Soc. Am. **42**, 191–198.
Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). ''Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey,'' J. Neurophysiol. **30**, 769–793.
Schouten, J. F. (**1940**). ''The residue and the mechanism of hearing,'' Proc. K. Ned. Akad. Wet. **43**, 991–999.
Schröder, M. R. (**1970**). ''Synthesis of low-peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **IT-16**, 85–89.
Shackleton, T. M., and Carlyon, R. P. (**1994**). ''The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination,'' J. Acoust. Soc. Am. **95**, 3529–3540.
Smoorenburg, G. F. (**1970**). ''Pitch perception of two-frequency stimuli,'' J. Acoust. Soc. Am. **48**, 924–942.
Srulovicz, P., and Goldstein, J. L. (**1983**). ''A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum,'' J. Acoust. Soc. Am. **73**, 1266–1276.
Terhardt, E. (**1974**). ''Pitch, consonance, and harmony,'' J. Acoust. Soc. Am. **55**, 1061–1069.
Wetherill, G. B. (**1963**). ''Sequential estimation of quantal response curves,'' J. R. Stat. Soc. Ser. B. Methodol. **B25**, 1–48.
Wichmann, F. A., and Hill, N. J. (**2001a**). ''The psychometric function. I. Fitting, sampling, and goodness of fit,'' Percept. Psychophys. **63**, 1293–1313.
Wichmann, F. A., and Hill, N. J. (**2001b**). ''The psychometric function. II. Bootstrap-based confidence intervals and sampling,'' Percept. Psychophys. **63**, 1314–1329.
Wightman, F. L. (**1973a**). ''Pitch and stimulus fine structure,'' J. Acoust. Soc. Am. **54**, 397–406.
Wightman, F. L. (**1973b**). ''The pattern-transformation model of pitch,'' J. Acoust. Soc. Am. **54**, 407–416.

Normal vibration frequencies of the vocal ligament

Ingo R. Titze[a)]
Department of Speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa 52242 and National Center for Voice and Speech, The Denver Center for the Performing Arts, Denver, Colorado 80204

Eric J. Hunter
National Center for Voice and Speech, The Denver Center for the Performing Arts, The University of Iowa, Iowa City, Iowa 52242

(Received 16 May 2003; revised 18 January 2004; accepted 14 February 2004)

The vocal ligament is the tension-bearing element in the vocal folds at high pitches. It has traditionally been treated as a vibrating string, with only length and longitudinal stress governing its normal mode frequencies. Results of this investigation show that, when bending stiffness and variable cross section are included, the lowest normal mode frequency can more than double, depending on the strain of the ligament. This suggests that much higher phonation frequencies may be achievable than heretofore thought for a given vocal fold length (e.g., nearly 1000 Hz at 50% elongation over cadaveric resting length). It also brings back into the discussion the concept of ''damping,'' an old misnomer for a reduction of the effective length of vibration of the vocal folds by relatively stiff boundary segments known as macula flavae. A formula is given for correcting the ideal string equation for the lowest mode frequency to include bending stiffness and macula flavae effects. [DOI: 10.1121/1.1698832]

PACS numbers: 43.70.Aj [AL] Pages: 2264–2269

I. INTRODUCTION

The vocal ligament is a portion of the nonmuscular layers of the vocal fold known as the lamina propria (layers in motion). It comprises the intermediate and deep layers of the lamina propria (Hirano, 1975). The ligament is composed of densely packed collagen and elastin fibers, which are aligned in a nearly parallel fashion and course anterio-posteriorally between the vocal process of the arytenoid cartilage and the thyroid cartilage (Fig. 1).

From a functional point of view, the vocal ligament limits its vocal fold elongation and helps position the vocal fold when the arytenoid cartilage moves. More importantly, for phonation, the vocal ligament supports large tensile stresses for high pitched sounds in singing, as well as for squeals and falsetto productions in spontaneous vocalizations. Because the ligament is thin (a few square mm in cross-section), large tensile stresses can be obtained with moderate muscle forces produced by the intrinsic laryngeal muscles (primarily the cricothyroid muscle). Among mammalian species, some have a highly developed vocal ligament (like the pig) while others have a poorly developed or nonexistent vocal ligament (like the dog). For human phonation, van den Berg (1958) hypothesized that the vocal ligament was critical for vocal registers (falsetto versus chest) and for high-pitched singing.

In previous studies of pitch control (e.g., Titze *et al.*, 1989), our assumption has always been that the ligament vibrates as a string with fixed boundary conditions, a uniform cross-section, uniform tension, and negligible bending stiffness. With these assumptions, the normal mode frequencies were easy to predict using the ideal string law. But recent attempts to reconcile human phonation frequencies with biomechanical data on vocal fold length and stress-strain characteristics of the ligament suggest that the ideal string law may underestimate the normal mode frequencies. Other authors (Descout *et al.*, 1980; Perrier, 1982; Perrier *et al.*, 1982; Guérin, 1983) recognized the ''beam'' nature of the vocal folds early on, predicting some natural frequencies based on bending stiffness. These predictions were not based on measured stress-strain curves of the vocal ligament, but rather on an average human tissue Young's modulus. Guérin also included large-amplitude vibration, for which the natural frequencies were dependent on amplitude of vibration and, hence, lung pressure. Bickley (1987) and Bickley and Brown (1987) claimed that the bending beam model is a better predictor of young children's F_0 in speech than either the ideal string model or the mass-spring model, the former predicting values too high and the latter too low. They stated that nonuniform tissue growth in early development may determine which model fits the anatomy best at any given age.

Forty years ago, Fletcher (1964) had a similar concern about the normal mode frequencies of a stiff piano string. It was known that the ideal string law

$$F = \frac{1}{2L}\sqrt{\frac{\sigma}{\rho}}, \tag{1}$$

where F is the natural frequency of the lowest mode, L is the length of the string, σ is the longitudinal stress, and ρ is the material density, underestimated the vibration frequency of a stiff piano string. The error was small, but when the proper bending stiffness was included in the analysis, the corrections matched with observations. In the case of the vocal ligament, the effect of including bending stiffness (and, in

a)Electronic mail: ingo-titze@uiowa.edu

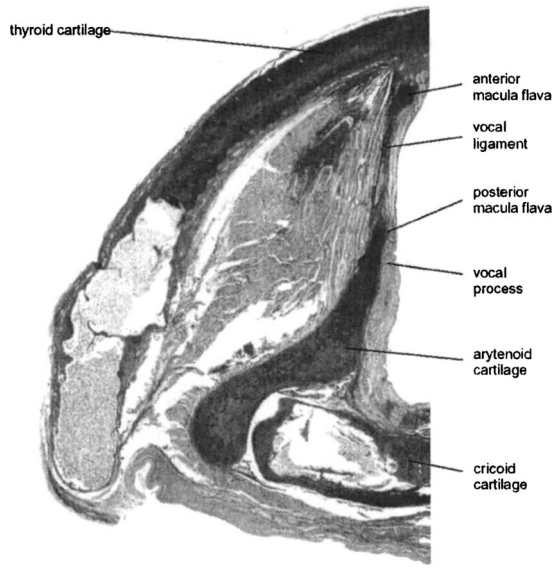


FIG. 1. Histological section through a male adult vocal fold showing vocal ligament in dark stain (after Hirano and Sato, 1993).

addition, variable cross-section) may be more than a small correction because the endpoints of the vocal ligament (the macula flavae) widen the ''string,'' possibly causing the boundary conditions to change more dramatically.

The purpose of this paper is to reanalyze the normal mode frequencies of the vocal ligament with the inclusion of bending stiffness and variable cross-section (at the macula flavae) in addition to tensile restoring forces. Both analytical and finite-element computational methods will be employed. The result will be based on actual measurements performed earlier on human vocal ligaments (Min *et al.*, 1995).

II. REVIEW OF ANALYTICAL TREATMENT OF LIGAMENT VIBRATION

A theoretical analysis of vibration of beams with longitudinal tension and bending stiffness was given by Morse (1936, reprinted 1976, pp. 166–170). The equation of motion is fourth order,

$$\rho\frac{\partial^2\xi}{\partial t^2}+R\frac{\partial\xi}{\partial t}=\sigma\frac{\partial^2\xi}{\partial x^2}-E\kappa^2\frac{\partial^4\xi}{\partial x^4}, \tag{2}$$

where ξ is the transverse displacement, ρ is the material density, R is the viscous loss coefficient (force per unit volume per m/s velocity), σ is the longitudinal stress (tension per unit area), E is the Young's modulus, and κ is the radius of gyration (for bending moments in the tissue). The loss term (second on the left side) was not included in Morse's discussion, but was included by Fletcher (1964) in the treatment of a stiff piano string. For a steel wire under high tension, the second term on the right side is small in comparison to the first, but in vocal fold tissue the two terms appear to be of comparable size, as will be seen.

For normal modes, for which the loss term R is neglected, a partial wave solution of the form

$$\xi(x,t)=Ae^{i\omega t-\gamma x} \tag{3}$$

satisfies the equations. Substitution of Eq. (3) into Eq. (2) yields the characteristic equation for wave propagation,

$$-\rho\omega^2=\sigma\gamma^2-E\kappa^2\gamma^4, \tag{4}$$

which has four algebraic solutions for the propagation constant γ,

$$\gamma=\pm\left\{\frac{\sigma}{2E\kappa^2}\pm\left[\left(\frac{\sigma}{2E\kappa^2}\right)^2+\frac{\rho\omega^2}{E\kappa^2}\right]^{1/2}\right\}^{1/2}. \tag{5}$$

For a ligament of rectangular cross section, the radius of gyration κ is defined as

$$\kappa=w/\sqrt{12}, \tag{6}$$

where w is the width of the ligament in the direction of transverse vibration (Morse, 1976, p. 153; Gieck and Gieck, 1997). The relative sizes of the terms in Eq. (5) will be discussed later.

III. CLAMPED BOUNDARY CONDITIONS

We assume clamped boundary conditions at the posterior $(x=0)$ and the anterior $(x=L)$ endpoints. This implies that both the displacement $\xi(x)$ and the derivative $\partial\xi/\partial x$ need to vanish at these boundaries,

$$\xi(0,t)=\xi(L,t)=0, \tag{7}$$

$$\frac{\partial\xi}{\partial x}(0,t)=\frac{\partial\xi}{\partial x}(L,t)=0. \tag{8}$$

With these boundary conditions, a complete solution to the fourth-order differential equation is written according to Morse (1936) or Fletcher (1964) as

$$\xi(x,t)=[A_1\cos\gamma_1x+B_1\sin\gamma_1x+A_2\cosh\gamma_2x+B_2\sinh\gamma_2x]e^{i\omega t}, \tag{9}$$

where

$$\gamma_1=-i\left\{\frac{\sigma}{2E\kappa^2}-\left[\left(\frac{\sigma}{2E\kappa^2}\right)^2+\frac{\rho\omega^2}{E\kappa^2}\right]^{1/2}\right\}^{1/2}, \tag{10}$$

$$\gamma_2=\left(\frac{\sigma}{E\kappa^2}+\gamma_1^2\right)^{1/2} \tag{11}$$

from Eq. (5). Both γ_1 and γ_2 are written to be positive and real, their imaginary and negative counterparts in Eq. (5) having been taken into consideration by choosing both trigonometric and hyperbolic functions in Eq. (9). Furthermore, since all four partial functions in brackets in Eq. (9) are linearly independent of each other, and since four arbitrary constants have been introduced, the solution is guaranteed to be complete and able to accommodate the four boundary conditions [Eqs. (7) and (8)].

Solution of the above equations for the constants A_1, B_1, A_2, and B_2 is a matter of algebra. Substituting Eq. (9) into Eqs. (7) and (8) yields four equations in four unknowns. The results are written here for ease of sequential numerical solution. Assuming A_1 to be an overall scale factor,

$$A_2=-A_1, \tag{12}$$

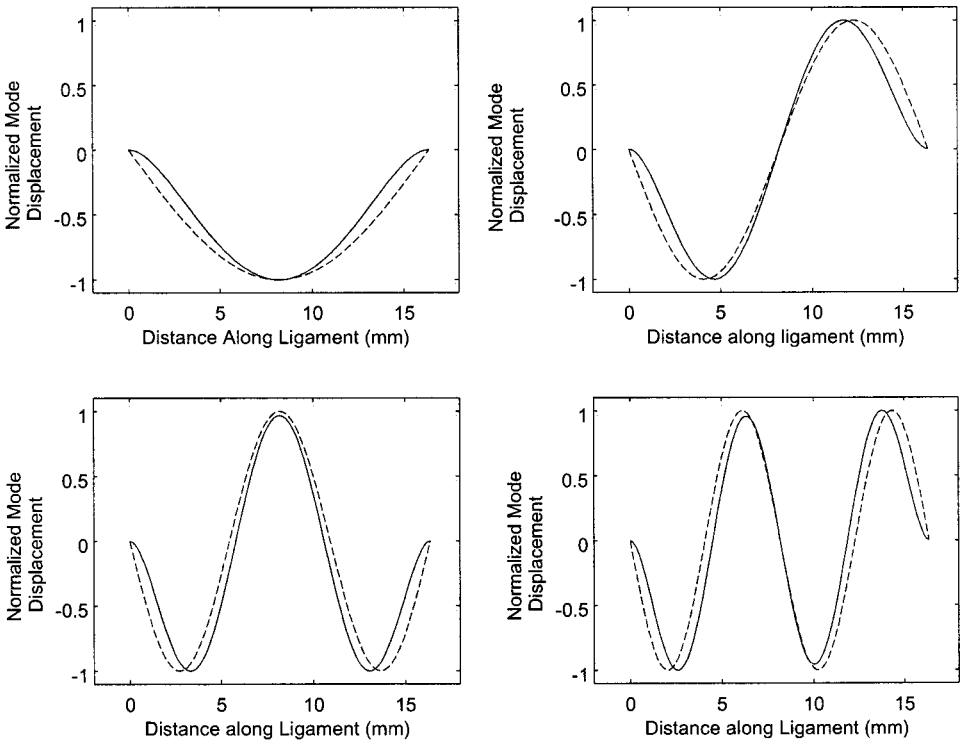


FIG. 2. Mode patterns for a stiff ligament (solid lines) and a classical "string" ligament (dashed lines) for the first four modes.

$$B_1 = \frac{-\gamma_1 \sin \gamma_1 L - \gamma_2 \sinh \gamma_2 L}{\gamma_1 \cosh \gamma_2 L - \gamma_1 \cos \gamma_1 L} A_1, \tag{13}$$

$$B_2 = -(\gamma_1/\gamma_2) B_1, \tag{14}$$

$$2\gamma_1\gamma_2(\cos \gamma_1 L \cosh \gamma_2 L - 1) = (\gamma_2^2 - \gamma_1^2)\sin \gamma_1 L \sinh \gamma_2 L. \tag{15}$$

A_1 cannot be determined because the four algebraic equations [Eqs. (7) and (8)] are homogeneous, as they always are for normal mode solutions. Instead, a condition is imposed on the propagation constants γ_1 and γ_2 via Eq. (15). Unfortunately, this equation is transcendental and can only be solved numerically to a high degree of accuracy. For this purpose, we assume that

$$\gamma_1 = \gamma_{1n} = \frac{n\pi}{L} + \delta_n, \tag{16}$$

where n is a mode integer, $n\pi/L$ is the wave number for a classical (nonbending) string, and δ_n is a correction factor. Equation (15) can be solved by computer, with repeated trials of δ_n, realizing that γ_2 is related to γ_1 by Eq. (11). Numerical results will be given below.

As a final step in the boundary value problem, the normal mode frequencies are calculated from Eq. (10) by solving for ω,

$$\omega_n = \left(\frac{\sigma}{\rho}\right)^{1/2}\left[\gamma_{1n}^2 + \frac{E\kappa^2}{\sigma\rho}\gamma_{1n}^4\right]^{1/2}. \tag{17}$$

If the material were linearly elastic under longitudinal tension, then

$$\sigma = \varepsilon E, \tag{18}$$

and E/σ in Eq. (17) could be replaced by the inverse of the longitudinal strain, $1/\varepsilon$. It is readily seen, then, that the bending stiffness term (second term in brackets) is inversely proportional to ε. The lower the strain, the more the bending stiffness will dominate relative to the tension term.

By substituting Eq. (16) into Eq. (17), making a first-order expansion of the square root term, and retaining only first power terms of δ_n, Morse (1936) approximated the normal mode frequencies for a beam under tension to be

$$\omega_n \approx \frac{n\pi}{L}\left(\frac{\sigma}{\rho}\right)^{1/2}\left[1 + \frac{2}{\pi}B^{1/2} + \left(\frac{4}{\pi^2} + \frac{n^2}{2}\right)B\right], \tag{19}$$

where

$$B = \frac{\pi^2 E \kappa^2}{\sigma L^2} = \frac{\pi^2 \kappa^2}{\varepsilon L^2}. \tag{20}$$

This approximation will be compared to the numerical solution for ω_n. Note that the value of B decreases with increasing strain.

Figure 2 shows the displacement function $\xi(x)$ for the first four modes. The mode amplitudes are all normalized to show the detail of the derivative at the endpoints. With bending stiffness (solid lines), the derivative goes to zero at $x = 0$ and $x = L$, whereas it retains a finite value for the classical string. Nodes and antinodes are shifted toward the center unless they occur at the center. In general, the modal patterns are not dramatically different. This is perhaps the reason why bending stiffness has generally been neglected; the differences do not show up on stroboscopic viewing.

Figure 3 shows the normal mode frequencies $\omega_n/2\pi$ plotted as a function of mode number n for several conditions. The following constants were chosen:

$$E = 91.7 \text{ kPa}, \quad \rho = 1040 \text{ kg/m}^3, \quad \sigma = 18.3 \text{ kPa},$$

$$L = 0.0136 \text{ m}.$$

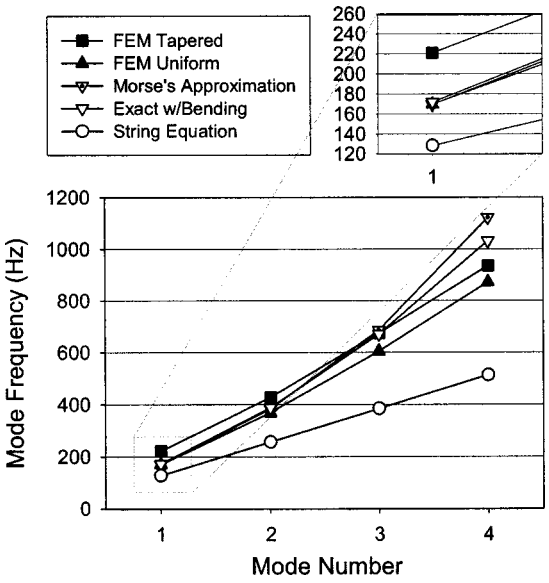


FIG. 3. Normal mode frequencies plotted against mode number (20% strain, $E=91.7$ kPa).

These constants were selected from data by Min *et al.* (1995) and our own anatomical measures from Hirano and Sato's (1993) histology. They would represent a medium-pitched phonatory adjustment for speech (20% elongation) if the ligament were the only tissue in vibration. Three of the curves in Fig. 3 are for solutions given so far. They are represented by the unfilled data symbols. Note that the classical string model (lowest curve—open circles) greatly underestimates the normal mode frequencies for a string with bending stiffness (exact solution in open diamonds). For mode 1, neglect of bending stiffness lowers the frequency from 170 to 130 Hz (see expanded view in upper right corner). For mode 4, neglect of bending stiffness lowers the frequency from 1030 to 500 Hz. Morse's approximation is shown with dotted diamonds. This approximation [Eq. (19)] is a good match to the exact solution up to mode 3. At mode 4, an 11% error is seen.

IV. VARIABLE CROSS-SECTION (MACULA FLAVAE)

A second step in the development of normal mode frequencies of the vocal ligament was the inclusion of the macula flavae, which are gradual widenings of the cross-sectional areas at the anterior and posterior endpoints (recall Fig. 1). It is believed that these macula flavae are nature's way of reinforcing the ligament at the boundaries in order to accommodate the vibrational stresses (both bending and shear). Using histological sections published by Hirano and Sato (1993) and Sato *et al.* (2003), an average resting length of 13.6 mm, and a cross section of 7.54 mm^2 as reported by Min *et al.* (1995), the geometry of the ligament was simplified to two cases with identical volume, as shown in Fig. 4: a rectangular parallelepiped (solid lines) and a parallelepiped with trapezoidal ends (shaded area). These geometries were implemented with 100 finite elements using ANSYS for a structural beam (elements LINK10 and BEAM54), where the cross section at any point was a perfect square.

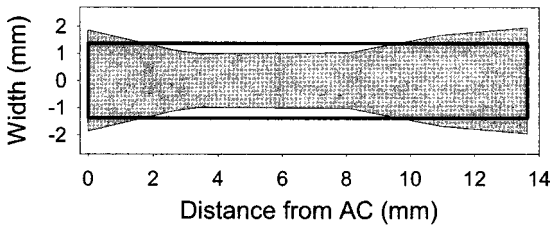


FIG. 4. Finite-element shapes of the ligament, uniform beam (solid lines) and a tapered beam with macula flavae (shaded area). Both have identical volume.

The uniform geometry incorporated both bending stiffness (BEAM 54) and no bending stiffness (LINK10), but ANSYS did not provide a nonbending option for the trapezoidal geometry (BEAM54). One finite element case (uniform geometry) could thus be checked against both the string equation and the exact analytical solution, and the second case was used to extend the theory beyond the analytical solution.

Returning to Fig. 3, the finite element results for modal frequencies are shown together with the analytical results (again, 20% strain, as in the analytical solution, with $E=91.7$ kPa). The uniform cross section with no bending stiffness was indistinguishable from the string equation; hence, no new data points are shown. The FEM solution with bending stiffness (closed diamonds) matched the analytical calculations exactly only for the lowest mode (open triangles). This match is also seen better in a zoomed-in version of the graph in the upper right. There is a discrepancy, however, for the higher modes. At modes 3 and 4, the finite element solution predicts lower frequencies than the exact analytical solution. This difference is attributed to the nature of the beam element equations when higher order shear effects are neglected. In other words, for a thick string with sizable cross section, bending and shear will generally interact, requiring another degree of freedom. But shear was not part of the overall analysis here because the differential equation was assumed to be one-dimensional [Eq. (2)]. Hence, an inconsistency exists in the FE analysis for higher modes where shear is set to zero but the cross section remains finite.

By adding the macula flavae, an additional increase in the normal mode frequencies (on the order of 30% at the lowest mode) is observed at this ligament length (tapered area—closed squares). For higher modes, the percent difference is diminished, but the absolute difference remains about the same.

V. NONLINEAR ELASTIC PROPERTIES OF THE VOCAL LIGAMENT

Min *et al.* (1995) quantified the nonlinear elastic properties of the human vocal ligament by mathematically fitting the stretching portion of a sinusoidal stress-release cycle with an exponential function:

$$\sigma = A(e^{B\varepsilon} - 1). \tag{21}$$

The A and B parameters were reported for each of eight ligaments with an average of $A=1.4$ kPa and $B=9.6$.

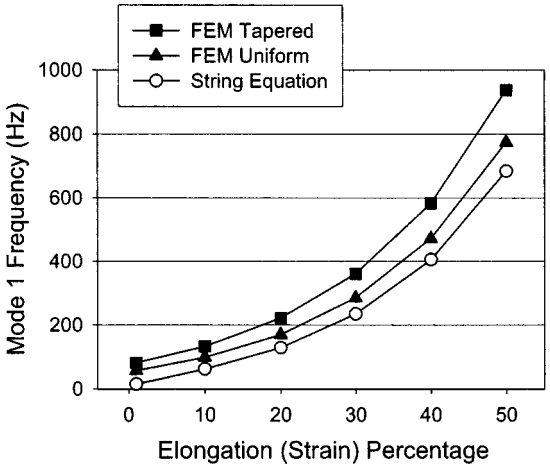


FIG. 5. Changes in the lowest normal mode frequency as a function of ligament strain, based on measured nonlinear stress-strain characteristics of human ligaments.

In the finite element solutions described above, four strains (1%, 10%, 20%, 30 and 40%) were used to solve for σ, E, and the first $(n=1)$ modal frequency. The tangent Young's modulus $\partial\sigma/\partial\varepsilon$ was used instead of the secant Young's modulus σ/ε because it is more appropriate when the stress-strain curve is nonlinear and there is a bias strain around which vibration takes place. Figure 5 shows how the lowest normal mode frequency changes with increasing ligament strain. The data points for 20% elongation are identical to those of Fig. 3 for mode 1. For 1% elongation, the contribution of bending stiffness and macula flavae raises the frequency from about 10 to 100 Hz, a profound difference. Zero strain would imply a zero string mode frequency because the tension would vanish. The Young's modulus, however, can remain finite due to nonfibrous (ground substance) tissue components. At 50% elongation, the inclusion of bending stiffness and macula flavae raises the mode 1 frequency from 680 to 770 Hz. The addition of the macula flavae (closed squares) raises the frequency by another 110 Hz. On a percentage basis, the bending stiffness contribution reduces with increasing strain, whereas the macula flavae contribution remains roughly constant (about 20%–30%).

VI. DISCUSSION AND CONCLUSIONS

Bending stiffness and nonuniform cross-sectional area of the vocal ligament appear to have an impact on the normal mode frequencies of the vocal ligament. With the mode shapes appearing visually similar to those of a simple string, it is understandable that the ideal string law has been the standard model. At low strains, the ideal string law underestimates the lowest mode frequency by an order of magnitude if bending stiffness is excluded. Asymptotically, as strain goes to zero, bending moments alone can account for the restoring force of the tissue. In previous modeling (Titze and Story, 2002), a ground substance Young's modulus was used to account for this restoring force. At higher strains, bending moments become less significant as tension takes over. Macula flavae contributions remain relatively constant at about 30% increase in mode 1 frequency. Implementing the full mathematical complexity of bending stiffness and nonuniform area may not be suitable for simple models of vocal fold vibration. A rule is therefore adopted for the lowest string mode frequency between 10% and 50% elongation:

$$f=\frac{1}{2L}\left(\frac{\sigma}{\rho}\right)^{1/2}(1-0.45\ln(\varepsilon)), \tag{22}$$

which corrects for bending stiffness and the macula flavae. For higher modes it is necessary to go back to Morse's approximation, which is only slightly more complicated.

The above results bring into discussion the old concept of ''damping,'' a misnomer for the reduction of the effective length of vibration of the vocal folds. We quote from an often-used textbook in speech and hearing science (Zemlin, 1997, pp. 167–168):

> ''Farnsworth (1940), Brodnitz (1959), Pressman (1942), and Pressman and Keleman (1955) attribute falsetto to a similar mechanism. That is, when the vocal folds have been tensed and lengthened as much as possible, further increases in pitch must be accompanied by a different mechanism, namely *damping*. The posterior portions of the vocal folds, in the region of the vocal processes, are firmly approximated and do not enter into vibration. As a result, *the length of the vibrating glottis is shortened considerably*.'' (Italics added by current authors.)

If vibration is ''damped'' at the endpoints, the vocal fold is effectively shortened, as if pressing a finger near the endpoints of a violin or guitar string. Perhaps a better term would be ''extended clamping'' because of the universal usage of damping for energy dissipation. In part, the macula flavae may provide this extended clamping with the addition of tapered endpoints. Thus, it may not require much adductory ''pressing'' to shorten the effective length. In fact, it is generally understood that falsetto is anything but ''pressing.'' Adduction is usually thought to be rather loose.

For low pitches and in modal voice productions, the vocal fold length is mostly below the cadaveric resting length. Ligament tension does not play much of a role. The stiffness of the thyroarytenoid muscle fibers then dominate in F_0 control, rather than the ligament. It is not yet clear if bending stiffness plays a significant role in thyroarytenoid muscle vibration.

ACKNOWLEDGMENT

This work was supported by NIH Grant No. 1 R01 DC04347.

Bickley, C. (**1987**). ''Acoustic evidence for the development of speech,'' Cambridge, MA: unpublished Ph.D. dissertation, Massachusets Institute of Technology.

Bickley, C., and Brown, K. (**1987**). ''Bending-beam model of vocal-fold vibration,'' J. Acoust. Soc. Am. **82**, S16.

Brodnitz, F. S. (**1959**). *Vocal Rehabilitation* (Whiting Pressing, Rochester, MN).

Descout, R., Auloge, J. Y., and Guérin, B. (**1980**). ''Continuous model of the vocal source,'' I.C.A.S.S.P., pp. 61–64.

Farnsworth, D. W. (**1940**). ''High-speed motion pictures of the vocal cords,'' Bell Lab. Rec. **18**, 203–208.

Fletcher, H. (**1964**). ''Normal vibration frequencies of a stiff piano string,'' J. Acoust. Soc. Am. **36**, 203–209.

Gieck, K., and Gieck, R. (**1997**). *Engineering Formulas, 7th Ed.* (McGraw-Hill, New York).
Guérin, B. (**1983**). *Vocal Fold Physiology Biomechanics, Acoustics and Phonatory Control*, edited by I. R. Titze and R. C. Scherer, Chapter 38: Effects of the source-tract interaction using vocal fold models (Denver Center for the Performing Arts, Denver).
Hirano, M. (**1975**). ''Phonosurgery: basic and clinical investigations,'' Otologia (Fukuoka) **21**, 129–440.
Hirano, M., and Sato, K. (**1993**), *Histological Color Atlas of the Human Larynx* (Singular, San Diego).
Min, Y. B., Titze, I. R., and Alipour-Haghighi, F. (**1995**). ''Stress-strain response of the human vocal ligament,'' Ann. Otol. Rhinol. Laryngol. **104**(7), 563–569.
Morse, P. M. (**1936**, reprinted **1976**). *Vibration and Sound* (Acoustical Society of America, New York) (also published by McGraw-Hill, New York, 1948).
Perrier, P. (**1982**). ''Etude d-un modèle continu des cordes vocals sous forme de deux pouters bi-articulées,'' Premiéres simulations, doctoral dissertation, I.N.P. Grenoble.
Perrier, P., Guérin, B., and Auloge, J. Y. (**1982**). ''Simulation d-un modèle continu de la source vocal,'' F.A.S.E./D.A.G.A., 1059–1062.
Pressman, J. (**1942**). ''Physiology of the vocal cords in phonation and respiration,'' Arch. Otolaryngol. **35**, 355–398.
Pressman, J., and Keleman, G. (**1955**). ''Physiology of the Larynx,'' Physiol. Rev. **35**, 506–554.
Rubin, H., and Hirt, C. C. (**1960**). ''The falsetto. A high-speed cinematographic study,'' Laryngoscope **70**, 1305–1324.
Sato, K., Hirano, M., and Nakashima, T. (**2003**). ''3D structure of themacula flava in the human vocal fold,'' Acta Otolaryngol. **123**, 269–273.
Titze, I. R. (**1989**). ''On the relation between subglottal pressure and fundamental-frequency in phonation,'' J. Acoust. Soc. Am. **85**, 901–906.
Titze, I. R., and Story, B. H. (**2002**). ''Rules for controlling low-dimensional vocal fold models with muscle activation,'' J. Acoust. Soc. Am. **112**(3, Pt 1), 1064–1076.
Titze, I. R., Luschei, E. S., and Hirano, M. (**1989**). ''Role of the thyroarytenoid muscle in regulation of fundamental frequency,'' J. Voice **3**(3), 213–224.
van den Berg, J. (**1958**). ''Myoelastic-aerodynamic theory of voice production,'' J. Speech Hear. Res. **3**, 227–244.
Zemlin, W. (**1997**). *Speech and Hearing Science. Anatomy & Physiology* (Prentice Hall, Englewood Cliffs, NJ).

Nonlinear dynamic analysis of voices before and after surgical excision of vocal polyps

Yu Zhang[a] and Clancy McGilligan
Department of Surgery, Division of Otolaryngology Head and Neck Surgery, University of Wisconsin Medical School, Madison, Wisconsin 53792-7375

Liang Zhou
Eye, Ear, Nose and Throat Hospital, Fudan University Medical School, 83 Fen Yang Road, Shanghai 200031, P.R. China

Mark Vig and Jack J. Jiang[b]
Department of Surgery, Division of Otolaryngology Head and Neck Surgery, University of Wisconsin Medical School, Madison, Wisconsin 53792-7375

(Received 9 October 2003; revised 9 February 2004; accepted 20 February 2004)

Phase space reconstruction, correlation dimension, and second-order entropy, methods from nonlinear dynamics, are used to analyze sustained vowels generated by patients before and after surgical excision of vocal polyps. Two conventional acoustic perturbation parameters, jitter and shimmer, are also employed to analyze voices before and after surgery. Presurgical and postsurgical analyses of jitter, shimmer, correlation dimension, and second-order entropy are statistically compared. Correlation dimension and second-order entropy show a statistically significant decrease after surgery, indicating reduced complexity and higher predictability of postsurgical voice dynamics. There is not a significant postsurgical difference in shimmer, although jitter shows a significant postsurgical decrease. The results suggest that jitter and shimmer should be applied to analyze disordered voices with caution; however, nonlinear dynamic methods may be useful for analyzing abnormal vocal function and quantitatively evaluating the effects of surgical excision of vocal polyps. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1699392]

PACS numbers: 43.70.Aj, 43.70.Bk, 43.25.Rq [AL] Pages: 2270–2277

I. INTRODUCTION

Vocal polyps, which are thought to result from voice misuse and overuse, are one of the most common benign laryngeal lesions.[1] Typically, they occur as unilateral lesions in the lamina propria and are located at the junction of the anterior and middle thirds of the vocal folds. Polyps disrupt glottal closure and produce irregular vocal-fold movement, resulting symptomatically in breathiness and hoarseness. Standard treatment includes speech therapy and surgical excision of lesions.[2] Excision of mass lesions of the vocal fold reduces laryngeal abnormalities and improves vocal function. Postsurgical improvements in the voices of patients with benign laryngeal lesions have been widely reported.[3–10] Voice analysis methods for measurement of postsurgical progress would be useful in the clinic. Although perceptual evaluation of vocal function is simple, the inherent subjectivity of perceptual judgments makes them maybe unreliable and difficult to quantify.[11,12] In addition, perceptual evaluation may fail to detect subtle changes in vocal function. Thus, reliable, objective measures for quantative assessment of postsurgical voice progress and behavior are highly desirable.

Recent studies have applied acoustic perturbation analysis to objectively evaluate the effects of surgery on the voices of patients with vocal-fold lesions.[3,5,7–9] Frequency perturbation, commonly called jitter, is a measure of short-term (cycle-to-cycle) variation in the fundamental frequency of a voice signal. Amplitude perturbation, commonly called shimmer, is a measure of short-term (cycle-to-cycle) variation in the amplitude of a voice signal. Uloza[7] found that jitter and shimmer are useful for studying the effects of surgery on vocal function; however, Zeitels *et al.*[9] found that there was not a significant difference in jitter measurements after surgical excision of vocal fold lesions, although there was a significant postsurgical decrease in shimmer measurements. Methodological issues have been raised regarding jitter and shimmer because of the sensitivity of these two parameters to variations in recording systems,[13,14] analysis systems,[15,16] and extraction algorithms.[14,17] In addition, systems employed for these perturbation parameters cannot reliably analyze strongly aperiodic signals.[8,18,19] It has been suggested that jitter and shimmer measurements are reliable for voice analysis only when their values are less than 5%[17,18] and may be unreliable when large instabilities in voice waveforms are observed.[19] Since jitter and shimmer only represent reliable parameters for nearly periodic voice signals under small perturbation conditions, seeking complementary objective measures capable of analyzing aperiodic voices and assessing the effects of surgery is important.

Nonlinear dynamic methods, including general dimension (Hausdorff dimension, information dimension, correlation dimension, etc.), entropy (Kolmogrov entropy, second-

[a]Electronic mail: zhang@surgery.wisc.edu
[b]Author to whom correspondence should be addressed. Electronic mail: jiang@surgery.wisc.edu

order entropy, etc.), and Lyapunov exponents, enable us to quantitatively describe aperiodic and chaotic phenomena[20] and have recently received a great deal of interest in the field of voice study.[21–29] Behrman and Baken[24] applied correlation dimension to the analysis of electroglottographic data from healthy and pathologic subjects. The effects of nonstationarity, noise, and finite signal length on the calculation of the correlation dimension of electroglottographic data were investigated. Hertrich *et al.*[25] found that the electroglottographic signals of patients with Parkinson disease and cerebellar diseases had significantly higher fractal dimensions than the electroglottographic signals of normal subjects. Giovanni *et al.*[26] found that maximal Lyapunov exponents statistically differentiated between normal voices and pathologic voices from patients with unilateral laryngeal paralysis. Jiang and Zhang[27] showed that the correlation dimensions of pathologic voices from patients with vocal polyps were significantly higher than those from normal subjects. These studies show the potential value of applying nonlinear dynamic methods to analyze pathologic voices.

The purpose of this study is to evaluate the ability of nonlinear dynamic methods to assess the effects of surgical excision of vocal polyps. Sustained vowels of 19 patients with laryngeal polyps are analyzed before and after surgical excision of lesions using phase space reconstructions, correlation dimension, and second-order entropy. In addition, jitter and shimmer, two conventional perturbation parameters, are calculated for the vowels before and after surgery. Presurgical and postsurgical analyses of jitter, shimmer, correlation dimension, and second-order entropy are statistically compared.

II. METHODS

A. Subjects and procedure

Nineteen patients with laryngeal polyps (13 women and 6 men) participated in this study. The average age of the patients was 41.6 years old (standard deviation=7.4). Patients received evaluation before and after surgical excision of polyps. On average, they were evaluated 5.0 days before surgery and 12.3 days after surgery. Voice recordings were made in a double-walled, sound-attenuated room. Subjects were asked to sustain the vowel /a/ at a comfortable pitch and intensity, as steadily and as long as possible. For each subject, three to four separate voice samples were recorded before and after surgical excision of polyps. In this manner we obtained 66 presurgical voice samples and 64 postsurgical voice samples. Voice signals were digitized with a 12-bit analog-to-digital converter (National Instruments AT-MIO-E-2) at a sampling rate of $f_s=20$ kHz by running a data acquisition program in LabVIEW 4.0 (National Instruments).

B. Data analysis

For all sustained vowels $x(t_i)$ with $t_i=iT_s$, $T_s=1/f_s$, $i=1,2,\ldots,N$, middle stationary-appearing segments with length 0.5 s ($N=10\,000$) were chosen for analysis. The dynamics of each voice segment were reconstructed in a phase space, which was then used to calculate correlation dimension and second-order entropy. A reconstructed phase space can be created by plotting a voice signal against itself at some time delay.[30] After an initial transient, a trajectory in phase space asymptotically approaches an invariant set, termed an attractor. The reconstructed phase space shows the dynamic behavior of a signal: a periodic signal produces a closed trajectory, while an aperiodic signal looks irregular and chaotic. Correlation dimension and second-order entropy are useful for their ability to describe irregular phenomena as well as their rapid convergence in numerical calculation.[31,32] Correlation dimension D_2 specifies the number of degrees of freedom needed to describe a system; a more complex system has a higher dimension, which means that more degrees of freedom may be needed to describe its dynamic state.[31] D_2 allows us to distinguish between deterministic chaos and random noise. Kolmogorov entropy quantifies the rate of loss of information about the state of a dynamic system as it evolves. Second-order entropy K_2 is the lower bound of Kolmogorov entropy and $K_2>0$ provides a sufficient condition for chaos.[32] Detailed descriptions concerning applications of phase space reconstruction, correlation dimension, and second-order entropy to phonation can be widely found in literatures.[21,27–29,33–36] In this study, our calculations were based on the numerical algorithms that we applied to study excised larynx phonations.[29] Briefly, an m-dimensional delay-coordinate phase space $\mathbf{X}_i=\{x(t_i),x(t_i-\tau),\ldots,x(t_i-(m-1)\tau)\}$ was reconstructed using the time delay technique,[30] where m is the embedding dimension and τ the time delay. m was determined according to the embedding theorem:[37] when $m>2D+1$ (D is the Hausdorff dimension), the reconstructed phase space is topologically equivalent to the original phase space. The proper time delay τ was estimated using the mutual information method proposed by Fraser and Swinney.[38] The improved algorithm proposed by Theiler[39] was used to calculate the correlation integral $C(r)$, where r is the radius around $\mathbf{X}_i$. Correlation integral $C(r)$ measures the number of distances between points in the reconstructed phase space that are smaller than the radius r. $C(r)$ has a power law behavior $C(r)\propto r^{D_2}e^{-m\tau K_2}$, which reveals the geometrical scaling property of the attractor.[32] Based on $C(r)$, D_2 and K_2 were manually estimated in the scaling region of the radius r with the increase of the embedding dimension m. The correlation dimension and its standard deviation were derived using a curve fit to the curve of $\log_2 C(r)$ vs. $\log_2 r$ in the scaling region.

Percent jitter and percent shimmer were computed for voices before and after surgical excision of polyps. The percent jitter is a cycle-to-cycle frequency perturbation measure and the percent shimmer is a cycle-to-cycle amplitude perturbation measure. Jitter and shimmer measure the short-term irregularities of the pitch period and the peak-to-peak amplitude of a voice signal, respectively. The percent jitter and percent shimmer were estimated using the Multi-Dimensional Voice Program (MDVP, Kay Elemetrics) where they are respectively defined as

$$\frac{[1/(K-1)]\sum_{i=1}^{K-1}|T_0^{(i)}-T_0^{(i+1)}|}{(1/K)\sum_{i=1}^{K}T_0^{(i)}}$$

and

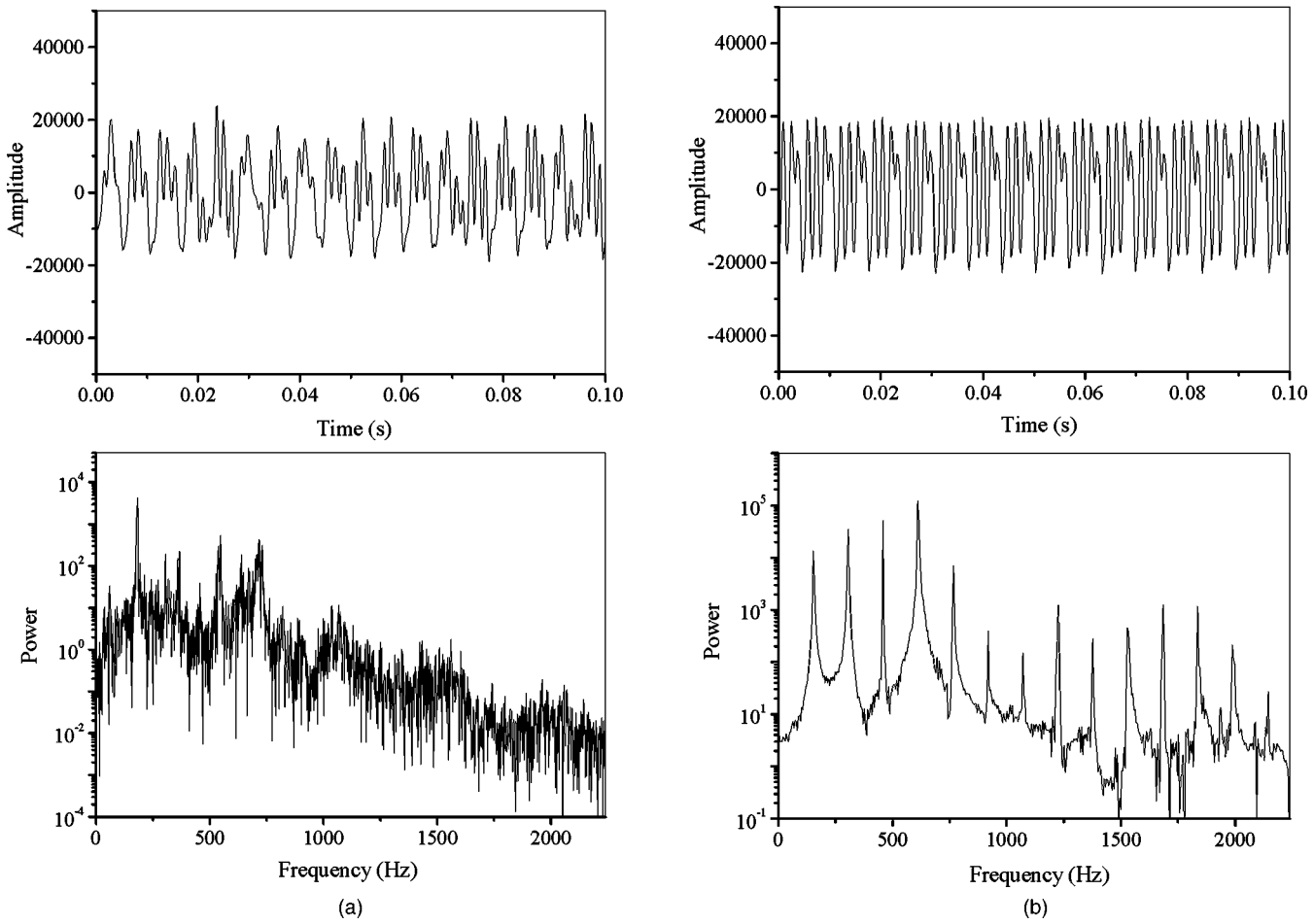


FIG. 1. (a) The 100-ms time series (the upper curve) and frequency spectrum (the lower curve) of the voice from a patient before surgery, where the voice length is 0.5 s and the frequency resolution is 1.22 Hz. (b) The 100-ms time series (the upper curve) and frequency spectrum (the lower curve) of the voice from the patient after surgery.

$$\frac{[1/(K-1)]\sum_{i=1}^{K-1}|A^{(i)}-A^{(i+1)}|}{(1/K)\sum_{i=1}^{K}A^{(i)}}, \quad (40)$$

$i=1,2,\ldots,K$, and K equals the number of pitch periods. $T_0^{(i)}$ and $A^{(i)}$ are the extracted pitch period and the peak-to-peak amplitude of a voice signal, respectively. MDVP can only reliably compute jitter and shimmer values for periodic and nearly-periodic signals; it cannot reliably analyze aperiodic signals.[40] Therefore, in this study, jitter and shimmer of nearly-periodic voices were computed to compare the presurgery and postsurgery groups.

C. Statistical analysis

For all voice samples recorded from 19 patients before and after surgery, correlation dimension D_2 and second-order entropy K_2 were calculated in their scaling regions in order to compare the results of presurgical and postsurgical analysis for each parameter. For nearly-periodic voices, jitter and shimmer were calculated using MDVP. Presurgical and postsurgical results were statistically compared. The Mann–Whitney rank sum test was employed using jitter, shimmer, D_2, and K_2 as dependent variables and the subject groups (presurgery and postsurgery) as independent variables. SigmaStat 2.0 (Jendel Scientific) software was used for statistical analysis.

III. RESULTS

A. The analyses of spectrum, jitter, and shimmer

Figure 1 shows the waveforms and spectra of presurgical and postsurgical voice samples from a single patient. Before surgery, the voice of the patient had the features of the type 3 signal defined by Titze,[18] namely, aperiodic waveform and noise-like broadband spectrum [see Fig. 1(a)]. However, after surgery, the voice had a nearly periodic waveform and strong discrete frequency peaks in the spectrum [see Fig. 1(b)], and thus could be classified as a Type 1 signal. The removal of polyps reduced noise components and eliminated voice hoarseness, which Cooper[10] also found. Postsurgical voices generally showed increased periodicity in this study. Jitter and shimmer of the voice after surgery in Fig. 1(b) were estimated as 1.97% and 1.82%, respectively. For all nearly periodic signals from 19 patients recorded before and after surgery, Figs. 2(a) and (b) show the distributions of the percent jitter and the percent shimmer for the presurgery and postsurgery groups, respectively. The results of statistical analysis of presurgical and postsurgical jitter and shimmer measurements are presented in Table I. Jitter decreased significantly after surgery ($p<0.001$). In contrast to this, shim-

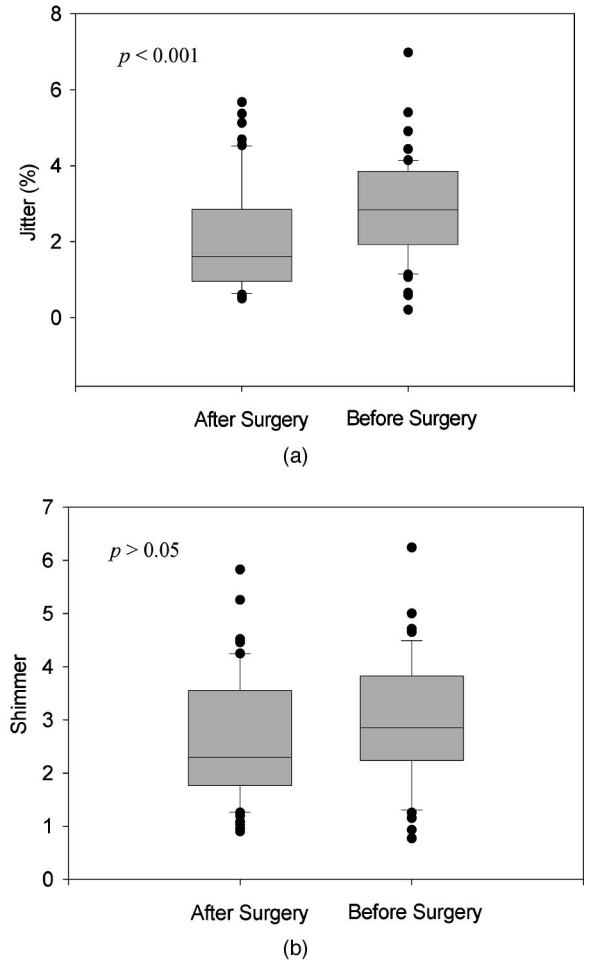


FIG. 2. The distributions of jitter and shimmer. (a) Jitter. (b) Shimmer, where the line inside the box marks the median and the dots are the outlying points.

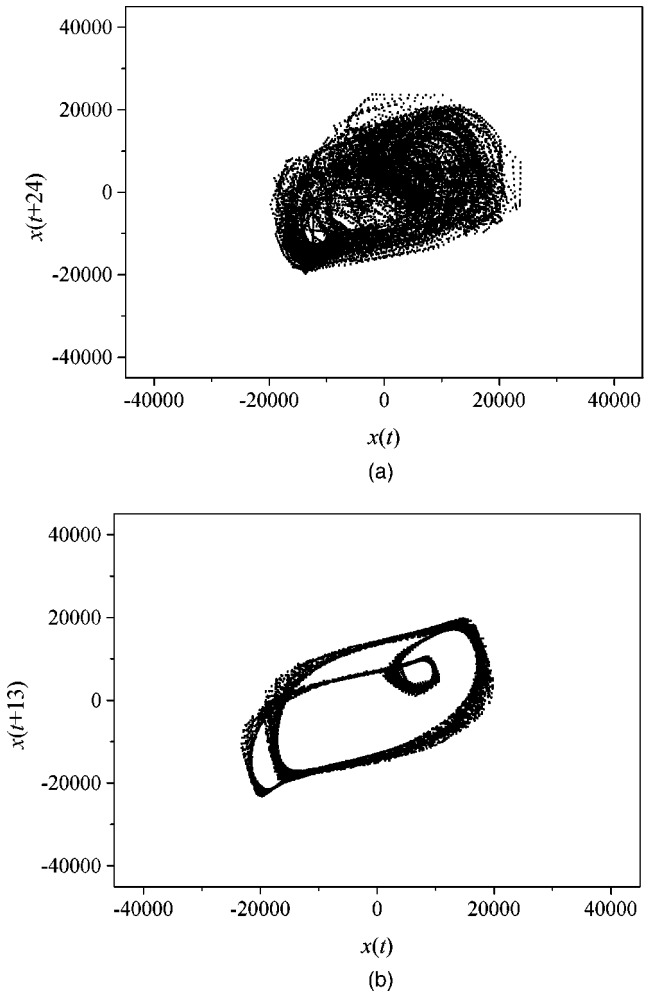


FIG. 3. Reconstructed $(x(t),x(t+\tau))$ phase spaces of the voices in Fig. 1. (a) Before surgery. (b) After surgery.

mer did not show a significant difference between the presurgery and postsurgery groups at the 0.05 significant confidence level.

B. Nonlinear dynamic analysis

Figures 3(a) and (b) show the trajectories in the reconstructed $(x(t),x(t+\tau))$ phase spaces of the presurgical and postsurgical voices of the patient from Fig. 1, where the lengths of the sustained vowels are 0.5 s ($N=10\,000$), and the proper time delay τ of the pre- and postsurgical voices were estimated as $24T_s$ and $13T_s$, respectively. Figures 4(a) and (b) show the graph of $\log_2 C(r)$ vs. $\log_2 r$ and its slope vs. r of the voice before surgery, respectively, in which the curves from top to bottom correspond to $m=1,2,\ldots,10$. If the radius r was too small, the effect of random noise was dominant so that the slopes of the $\log_2 C(r)$ vs. $\log_2 r$ curves did not converge with m. On the other hand, if the radius r was too large, the slopes approached zero. There existed a finite region ($8000<r<16\,000$) between these two regions, termed the scaling region, in which all the slopes of the $\log_2 C(r)$ vs. $\log_2 r$ curves converged. When m was significantly large, this gave the estimate of the correlation dimension $D_2=2.59\pm0.04$ of the voice before surgery [see Fig. 4(b)]. The correlation dimension of the voice after surgery was estimated in the same way, giving $D_2=1.17\pm0.02$ in the scaling region ($2200<r<9400$). Figure 4(c) shows the

TABLE I. Comparisons of presurgery and postsurgery groups for differences on jitter, shimmer, D_2, and K_2.

	Before surgery		After surgery		
	Median	25%–75% Range	Median	25%–75% Range	Mann–Whitney (p-value)
Jitter	2.84	1.93–3.85	1.61	1.0–2.82	<0.001
Shimmer	2.85	2.25–3.80	2.24	1.74–3.54	0.06[a]
D_2	2.01	1.32–2.53	1.48	1.24–1.86	<0.001
K_2	0.175	0.085–0.25	0.085	0.068–0.124	<0.001

[a]Not significant.

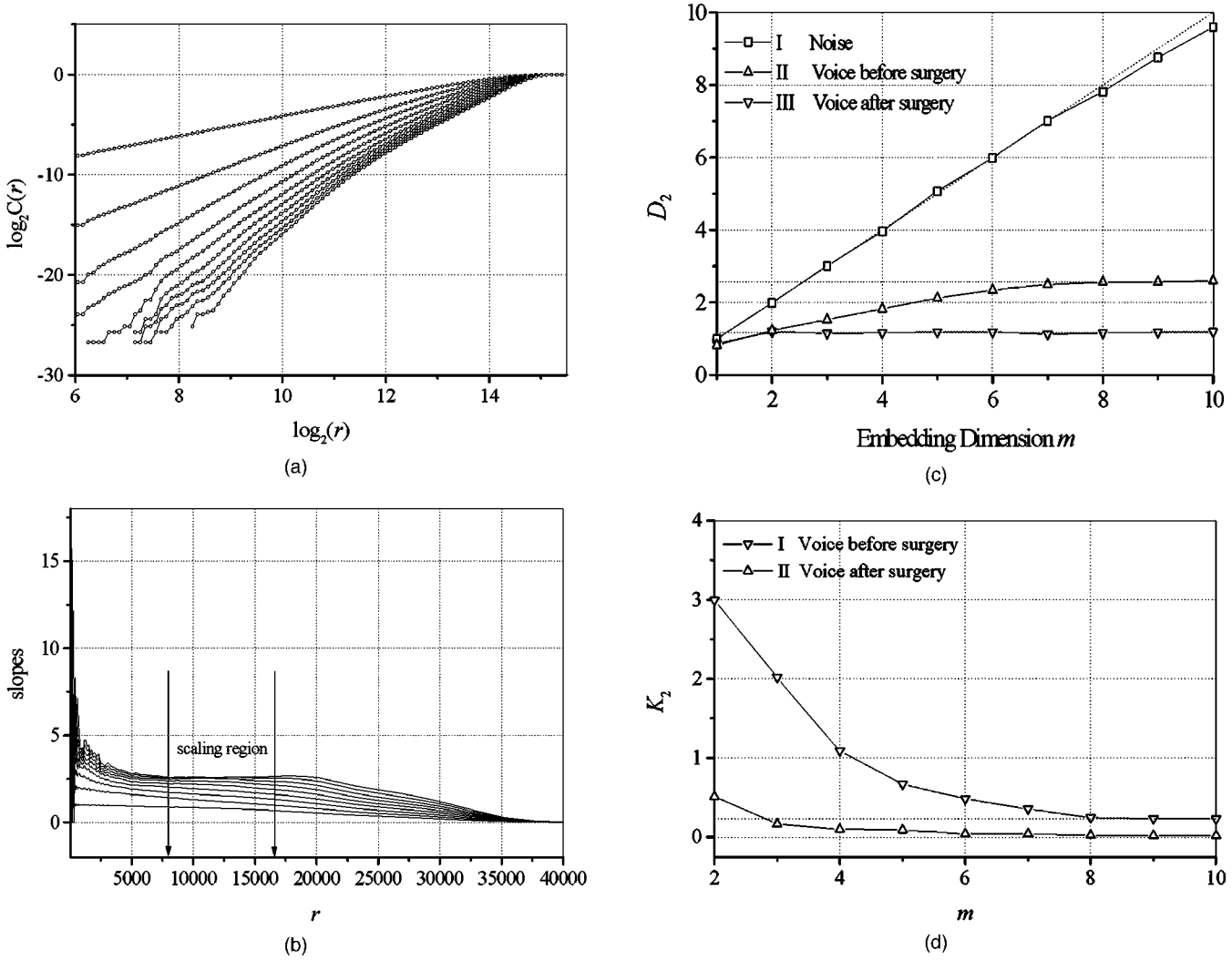


FIG. 4. (a) $\log_2 C(r)$ vs. $\log_2 r$ of the voice before surgery in Fig. 3(a), in which the curves from top to bottom correspond to $m=1,2,\ldots,10$, respectively. (b) The slope of Fig. 4(a) vs. r. (c) The dimension estimate versus m, in which curves I, II, and III correspond to random noise and the voice before and after surgery, respectively. (d) The entropy estimate versus m, in which curves I and II correspond to the voices before and after surgery.

relationship between the dimension estimate and m, in which curves I, II, and III correspond to random noise, the voice before surgery, and the voice after surgery, respectively. With an increase of m, the estimate of D_2 of random noise did not exhibit a saturation tendency; however, the estimated dimensions of the voices before and after surgery converged to 1.17 ± 0.02 and 2.59 ± 0.04, respectively. Figure 4(d) gives the dependence of the entropy estimate on m, in which curves I and II correspond to the voices before and after surgery, respectively. With the increase of m, the estimates of the second-order entropy K_2 of the voices before and after surgery converged to 0.24 ± 0.01 bits/T_s and 0.02 ± 0.01 bits/T_s, respectively.

For *all* the patients, Figs. 5(a) and (b) show the distributions of D_2 and K_2 for the presurgery and postsurgery groups. Results of statistical analysis of presurgical and postsurgical D_2 and K_2 measurements are displayed in Table I. In all of the subjects, the estimates of correlation dimensions were finite ($D_2<5$). Results of a Mann–Whitney rank sum test performed on D_2 and K_2 revealed that the correlation dimensions and second-order entropies of the group before surgery were both significantly higher than those of the group after surgery ($p<0.001$).

IV. DISCUSSION

We employed jitter and shimmer to analyze voices before and after surgical excision of vocal polyps. Recent studies have found that these perturbation parameters may be unsatisfactory since they are not reliable for aperiodic or chaotic voice signals.[8,17–19] It has been suggested that jitter and shimmer measurements are reliable for nearly periodic voice analysis.[17,18,40] Thus, in this study, we calculated jitter and shimmer of nearly periodic signals when statistically assessing the effects of vocal polyp excision. It was found that jitter values decreased significantly after surgery ($p<0.001$), but shimmer values did not show a significant decrease ($p>0.05$) (Fig. 2). On the other hand, Zeitels *et al.*[9] found that there was not a significant difference in jitter measurements after excision of benign vocal-fold lesions, although there was a significant decrease in shimmer measurements. The findings of our study, the study of Zeitels *et al.*[9]

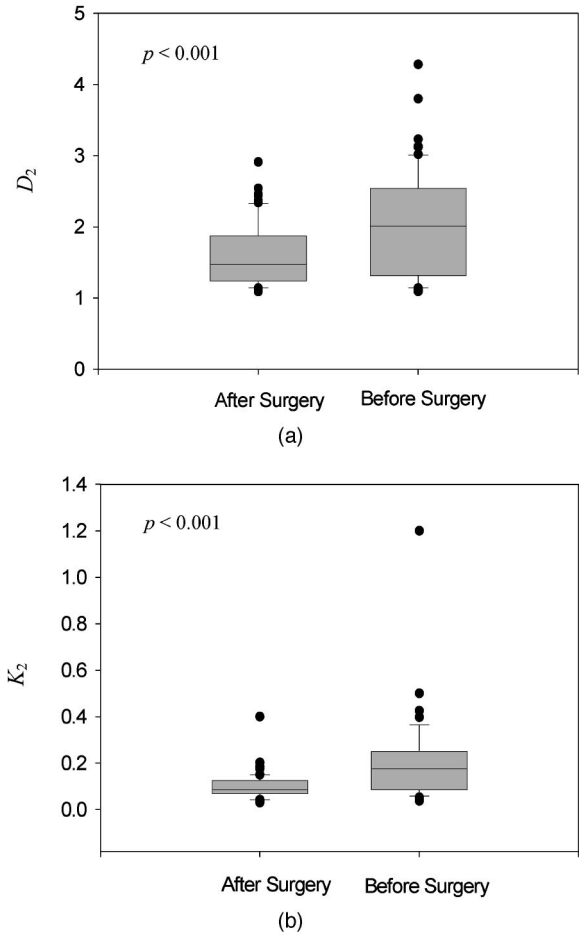


FIG. 5. The distribution of D_2 and K_2 of all the patients before and after surgery: (a) D_2 and (b) K_2, where the line inside the box marks the median and the dots are the outlying points.

and other studies suggest that these perturbation measures should be applied with caution, particularly when aperiodic or chaotic voice signals are being analyzed. A similar conclusion was arrived at via our recent experiments with excised larynges.[29] These findings are not surprising; jitter and shimmer are recognized as inadequate for analysis of far-from-periodic voices.

Nonlinear dynamic analysis has recently been applied to study disordered voices resulting from laryngeal pathologies. Titze *et al.*[21] found that the correlation dimension of a dysphonic subject with polyposis was higher than that of a healthy subject. Using the time series of F0 and amplitude derived from microphone signals, Kakita and Okamoto[23] found that the median correlation dimension value of 17 dysphonic subjects is higher than that of eight healthy subjects. In addition, fractal dimensions[25,27] and maximal Lyapunov exponents[26] have revealed statistically significant differences between normal voices and pathologic voices from patients with such laryngeal pathologies as vocal-fold paralysis, Parkinson's disease, cerebellar atrophy, and vocal polyps. Three types of voice signals have been statistically distinguished with nonlinear dynamic methods.[28] These studies show the potential value of applying nonlinear dynamic methods to the analysis of pathologic voices. The main purpose of our study was to evaluate the ability of nonlinear dynamic methods to analyze abnormal vocal function and quantify postsurgical voice progress. As far as we know, this is the first time nonlinear dynamic methods have been applied to evaluate the effects of surgical excision of vocal polyps. Compared with the irregular and chaotic structure of the voice before surgery, the trajectory of the voice after surgery was generally found to be regular and simple [see Figs. 3(a) and (b)]. Correlation dimension and second-order entropy were able to describe the aperiodic and irregular voice signal, as seen in Fig. 4. Correlation dimension and second-order entropy showed a statistically significant improvement ($p<0.001$) in postsurgical voices, as found in Fig. 5. Jitter and shimmer are objective parameters that can reliably measure the frequency and amplitude perturbations of nearly periodic voice signals. Our results suggest that nonlinear dynamic analysis might represent a complementary objective method that can measure periodic or aperiodic signals (not including random noise) and quantify postsurgical voice progress.

Noise, finite signal length, and nonstationarity might affect estimates of the correlation dimensions of phonation signals.[24,29,33,34] In this study, we attempted to ensure reliable estimates in the following manner. Voice recordings were made in a double-walled, sound-attenuated room in order to reduce the effects of external noise perturbations. Figures 4(a) and (b) show that there existed substantial scaling regions in which all the slopes of the $\log_2 C(r)$ vs. $\log_2 r$ curves converged, suggesting that noise perturbations did not affect the estimate of the correlation dimension made in the scaling region. In addition, the determined vowel length $N=10\,000$ satisfied the requirement proposed by Eckmann and Ruelle.[41] For the vowel length $N=10\,000$ (0.5 s), the upper bound of a reasonable estimate of the correlation dimension D_2 is about 8 according to the formula of Eckmann and Ruelle.[41] For all sustained vowels in this study, estimates of correlation dimension satisfied $D_2<5$. Thus, the length $N=10\,000$ was long enough for reasonable estimates of the correlation dimensions of sustained vowels. Furthermore, a system is considered stationary if system parameters have no obvious time variations. For the voice production system, stationarity requires stable phonation conditions, i.e., phonation parameters such as glottal configuration and subglottal pressure must lack obvious time variations. In this study, we attempted to ensure stable conditions of voice productions using the following procedures: we asked subjects to sustain the vowel /a/ at a comfortable pitch and intensity, as steadily and as long as possible, and then chose relatively short (0.5 s), stationary-appearing middle segments of the sustained vowels for analysis. Sustained vowel segments also have been previously applied for conventional acoustic analyses.[3,5,7–9] Short, middle segments of sustained vowels may not include information concerning transient effects, such as those occurring at the beginning and end of voice productions. Thus the analysis of short, middle portions of sustained vowel segments reduces the risk of nonstationarity. Using these approaches, we found that the scaling regions of sustained vowels before and after surgery were substantial and the slopes of $\log_2 C(r)$ vs. $\log_2 r$ could be estimated reliably [see Figs. 4(a) and (b)]. Previous studies[21,23,27–29,33–36] found that stationary voice signals had substantial scaling

regions and estimates of dimensions converged with the increase of the embedding dimension. Behrman and Baken[24] analyzed generally nonstationary electroglottographic data and found constricted scaling regions and nonconstant slopes when calculating correlation dimension. Electroglottographic signals often exhibit direct current baseline shift.[42,43] In addition, electroglottographic signals are significantly affected by electrode placement and contact, as well as head movement during phonation. These factors might possibly affect estimation of correlation dimensions by introducing noise and additional nonstationarity. Further discussions concerning the effects of these factors on estimating dimension of electroglottographic data are out of the scope of this paper.

The estimates of correlation dimensions and second-order entropy of voices before and after surgery converged with the increase of the embedding dimension, differing from nonconvergent values of random noise, as shown in Fig. 4(c). This indicates that the voices from patients with polyps had low-dimensional characteristics. Thus finite state variables might be sufficient to describe the dynamics of voices from patients with vocal polyps. This is in agreement with the finite element simulations of Jiang *et al.*,[44] where the vibrations of the vocal folds with a unilateral vocal mass were dominated by the first few vibratory modes. After vocal polyp excision, correlation dimension and second-order entropy showed a statistically significant decrease (Fig. 5). A postsurgical decrease in correlation dimension indicates that postsurgical phonations are less complex and may require fewer state variables to describe their dynamics. A postsurgical decrease in second-order entropy indicates the increased predictability and lower rate of loss of information about the dynamic state of postsurgical phonations. Both of these decreases imply postsurgical voice improvements. Thus nonlinear dynamic analysis successfully quantified postsurgical voice progress by showing decreased dynamic complexity and increased predictability of postsurgical phonations. This is in agreement with the findings of our recent study of a vocal fold model with a unilateral vocal polyp,[45] where the polyp induced complex, chaotic vocal fold vibrations, while a significant reduction in polyp size, e.g., resulting from surgical excision, caused simple, periodic oscillations.

V. CONCLUSION

In this paper, we applied phase space reconstructions, correlation dimension, and second-order entropy, methods from nonlinear dynamics, to analyze the sustained vowels of patients before and after surgical excision of vocal polyps. We also employed jitter and shimmer, two conventional acoustic parameters, to analyze the sustained vowels. The voices from patients showed a statistically significant postsurgical decrease in jitter, but not in shimmer. However, both correlation dimension and second-order entropy showed a statistically significant postsurgical decrease, leading us to conclude that they are capable of analyzing the voices of patients before and after surgery. Finite values of dimension and entropy were shown. Nonlinear dynamic methods have been widely applied to the analysis of pathologic human voices.[21–28] Our study suggests that correlation dimension and second-order entropy might be capable of quantitatively analyzing voices in order to assess the effects of surgery. Nonlinear dynamic analysis might be usefully applied to complement existing voice analysis methods in the clinic.

ACKNOWLEDGMENTS

This study was supported by NIH Grant No. 1-RO1DC006019-01 from the National Institute of Deafness and other Communication Disorders.

[1] F. Chagnon and R. E. Stone, ''Nodules and polyps,'' in *Organic Voice Disorders Assessment and Treatment*, edited by W. S. Brown, B. P. Vinson, and M. A. Crary (Singular, San Diego, 2000), p. 219.
[2] M. S. Strong and C. W. Vaughan, ''Vocal cord nodules and polyps-the role of surgical treatment,'' Laryngoscope **81**, 911–923 (1971).
[3] G. S. Berke, B. R. Gerratt, and D. G. Hanson, ''An acoustic analysis of the effects of surgical therapy on voice quality,'' Otolaryngol.-Head Neck Surg. **91**, 502–508 (1983).
[4] M. Bouchayer and G. Cornut, ''Microsurgical treatment of benign vocal fold lesions: indications, technique, results,'' Folia Phoniatr. **44**, 155–184 (1992).
[5] P. Woo, J. Casper, R. Colton, and D. Brewer, ''Aerodynamic and stroboscopic findings before and after microlaryngeal phonosurgery,'' J. Voice **8**, 186–194 (1994).
[6] R. W. Bastian, ''Vocal Fold Microsurgery in Singers,'' J. Voice **10**, 389–404 (1996).
[7] V. Uloza, ''Effects on voice by endolaryngeal microsurgery,'' Eur. Arch. Otorhinolaryngol. **256**, 312–315 (1999).
[8] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. VanDeHeyning, M. Remacle, V. Woisard, and Committee on Phoniatric of the European Laryngological Society (ELS), ''A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS),'' Eur. Arch. Otorhinolaryngol. **258**, 77–82 (2001).
[9] S. M. Zeitels, R. E. Hillman, R. Desloge, M. Mauri, and P. B. Doyle, ''Phonomicrosurgery in singers and performing artists: treatment outcomes, management theories, and future directions,'' Ann. Otol. Rhinol. Laryngol. **111**, 21–40 (2002).
[10] M. Cooper, ''Spectrographic analysis of fundamental frequency and hoarseness before and after vocal rehabilitation,'' J. Speech Hear. Res. **39**, 286–297 (1974).
[11] J. Kreiman and B. R. Gerratt, ''Validity of rating scale measures of voice quality,'' J. Acoust. Soc. Am. **104**, 1598–1608 (1998).
[12] B. R. Gerratt and J. Kreiman, ''Measuring vocal quality with speech synthesis,'' J. Acoust. Soc. Am. **110**, 2560–2566 (2001).
[13] J. J. Jiang, E. Lin, and D. G. Hanson, ''Effect of tape recording on perturbation measures,'' J. Speech Hear. Res. **41**, 1–31–1041 (1998).
[14] I. R. Titze, Y. Horii, and R. C. Scherer, ''Some technical considerations in voice perturbation measurements,'' J. Speech Hear. Res. **30**, 252–260 (1987).
[15] M. P. Karnell, R. S. Scherer, and L. B. Fischer, ''Comparison of fundamental frequency and perturbation measurements among three analysis systems,'' J. Voice **9**, 383 (1995).
[16] S. Bielamowicz, J. Kreiman, B. R. Gerratt, M. S. Dauer, and G. S. Berke, ''Comparision of voice analysis systems for perturbation measurement,'' J. Speech Hear. Res. **39**, 126–134 (1996).
[17] I. R. Titze and H. Liang, ''Comparison of F0 Extraction Method for High-Precision Voice Perturbation Measurements,'' J. Speech Hear. Res. **36**, 1120–1133 (1993).
[18] I. R. Titze, ''Workshop on acoustic voice analysis; Summary statement,'' National Center for Voice and Speech, Denver, CO (1995), p. 36.
[19] M. P. Karnell, A. Chang, A. Smith, and H. Hoffman, ''Impact of signal type of validity of voice perturbation measures,'' NCVS Status and Progress Report **11**, 91–94 (1997).
[20] E. Ott, T. Sauer, and J. A. Yorke. *Coping with chaos: analysis of chaotic data and the exploitation of chaotic systems* (Wiley, New York, 1994).
[21] I. R. Titze, R. Baken, and H. Herzel, ''Evidence of chaos in vocal fold vibration,'' in *Vocal Fold Physiology: New Frontier in Basic Science*, edited by I. R. Titze (Singular, San Diego, 1993), pp. 143–188.

[22] H. Herzel, D. Berry, I. R. Titze, and M. Saleh, ''Analysis of vocal disorders with methods from nonlinear dynamics,'' J. Speech Hear. Res. **37**, 1008–1019 (1994).
[23] Y. Kakita and H. Okamoto, ''Visualizing the characteristics of vocal fluctuation from the viewpoint of chaos: an attempt toward qualitative quantification,'' in *Vocal Fold Physiology—Voice Quality Control*, edited by O. Fujimura and M. Hirano (Singular, San Diego, 1995), pp. 235–248.
[24] A. Behrman and R. J. Baken, ''Correlation dimension of electroglottographic data from healthy and pathologic subjects,'' J. Acoust. Soc. Am. **102**, 2371–2379 (1997).
[25] I. Hertrich, W. Lutzenberger, S. Spieker, and H. Ackermann, ''Fractal dimension of sustained vowel productions in neurological dysphonias: an acoustic and electroglottographic analysis,'' J. Acoust. Soc. Am. **102**, 652–654 (1997).
[26] A. Giovanni, M. Ouaknine, and J-M Triglia, ''Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis'' J. Voice **13**, 341 (1998).
[27] J. J. Jiang and Y. Zhang, ''Nonlinear dynamic analysis of speech from pathologic subjects,'' Electron. Lett. **38**, 294–295 (2002).
[28] Y. Zhang and J. J. Jiang, ''Nonlinear dynamic analysis of signal typing of pathological human voices,'' Electron. Lett. **39**, 1021–1023 (2003).
[29] J. J. Jiang, Y. Zhang, and C. N. Ford, ''Nonlinear dynamic of phonations in excised larynx experiments,'' J. Acoust. Soc. Am. **114**, 2198–2205 (2003).
[30] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, ''Geometry from a time series,'' Phys. Rev. Lett. **45**, 712 (1980).
[31] P. Grassberger and I. Procaccia, ''Measuring the strangeness of strange attractors,'' Physica D **9**, 189–208 (1983).
[32] P. Grassberger and I. Procaccia, ''Estimation of the Kolmogorov entropy from a chaotic signal,'' Phys. Rev. A **28**, 2591–2593 (1983).
[33] S. S. Narayanan and A. A. Alwan, ''A nonlinear dynamical systems analysis of fricative consonants,'' J. Acoust. Soc. Am. **97**, 2511 (1995).
[34] A. Kumar and S. K. Mullick, ''Nonlinear dynamical analysis of speech,'' J. Acoust. Soc. Am. **100**, 615–629 (1996).
[35] M.-H. Lee, J.-N. Lee, and K.-S. Soh, ''Chaos in segments from korean traditional singing and western singing,'' J. Acoust. Soc. Am. **103**, 1175–1182 (1998).
[36] A. M. Chen and H. Leung, ''Equalization of speech and audio signals using a nonlinear dynamical approach,'' IEEE Trans. Speech Audio Process. **7**, 356–360 (1999).
[37] F. Takens, ''Detecting strange attractors in turbulence,'' in *Lecture Notes in Mathematics*, edited by D. A. Rand and B. S. Young (Springer-Verlag, Berlin, 1981), Vol. 898, pp. 366–381.
[38] A. M. Fraser and H. L. Swinney, ''Independent coordinates for strange attractors from mutual information,'' Phys. Rev. A **33**, 1134–1140 (1986).
[39] J. Theiler, ''Spurious dimension from correlation algorithms applied to limited time series data,'' Phys. Rev. A **34**, 2427–2432 (1986).
[40] Kay Elemetrics Corp, *Multi-Dimensional Voice Program: Software Instruction Manual*, Pine Brook, NJ (1993).
[41] J. P. Eckmann and D. Ruelle, ''Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems,'' Physica D **56**, 185–187 (1992).
[42] D. G. Childers and A. K. Krishnamurthy, ''A critical review of Electroglottography and laryngostroboscopy,'' Crit. Rev. Biomed. Eng. **12**, 131–161 (1985).
[43] I. R. Titze, Summary statement, in *Workshop on acoustic voice analysis*, edited by D. Wong (National Center for Voice and Speech, Iowa City, IA, 1995).
[44] J. J. Jiang, C. E. Diaz, and D. G. Hanson, ''Finite element modeling of vocal fold vibration in normal phonation and hyperfunctional dysphonia: implication for the pathogenesis of vocal nodules,'' Ann. Otol. Rhinol. Laryngol. **107**, 603–609 (1998).
[45] Y. Zhang and J. J. Jiang, ''Chaotic vibrations of a vocal-fold model with a unilateral polyp,'' J. Acoust. Soc. Am. **115**, 1266–1269 (2004).

Spectral integration of synchronous and asynchronous cues to consonant identification

Emily Buss,[a] Joseph W. Hall III, and John H. Grose
The Department of Otolaryngology/Head and Neck Surgery at The University of North Carolina School of Medicine, 130 Mason Farm Road, CB#7070, 1115 Bioinformatics Building, Chapel Hill, North Carolina 27599

(Received 20 February 2003; revised 30 January 2004; accepted 9 February 2004)

The goal of the present research was to determine how well observers utilize cues to consonant identification from different spectral regions that occur asynchronously as opposed to synchronously across frequency; such an ability would be useful for processing speech in the context of a spectro-temporally complex masker (e.g., competing speech). This was assessed by obtaining masked identification thresholds for VCV speech material, of the form /*a*/ C /*a*/, under various conditions of 10-Hz or 20-Hz square-wave amplitude modulation (AM). The speech tokens were filtered into 2, 4, 8, or 16 contiguous log-spaced frequency bands spanning 0.1 to 10 kHz. Bands were then modulated, with the pattern of that AM being either coherent across bands or 180° out of phase for adjacent bands. In the *out-of-phase* conditions the odd-numbered bands had coherent phase AM and the even bands had coherent phase AM, but the AM pattern across these two subsets of bands were out of phase. Results from these two conditions, along with further conditions employing only modulated even- or odd-numbered bands, allowed performance to be compared between stimuli characterized by synchronous and asynchronous cues. Results indicate that observers are able to utilize asynchronously presented cues to consonant identification efficiently across a range of conditions. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1691035]

PACS numbers: 43.71.Es, 43.66.Mk [KWG] Pages: 2278–2285

I. INTRODUCTION

Masked speech identification improves when the masker is amplitude modulated (AM) (e.g., Miller and Licklider, 1950; Carhart *et al.*, 1966; Dirks *et al.*, 1969; Gustafsson and Arlinger, 1994). It has been suggested that this improvement is based on the introduction of brief epochs of enhanced signal-to-noise ratio at the masker modulation minima, during which advantageous speech cues are present (e.g., Miller and Licklider, 1950). Despite the fact that these epochs are distributed over time, the auditory system is adept at piecing together the cues into an interpretable percept. It is not clear, however, what role temporal coincidence of these ''glimpses'' across frequency plays in this result. Whereas the fluctuations in AM noise are coherent across frequency, the fluctuations characterizing many other kinds of maskers (e.g., speech maskers or ''cafeteria noise'') are not. When speech is presented at a poor signal-to-noise ratio in such backgrounds, the observer may be momentarily presented with information in some spectral regions at a more favorable signal-to-noise ratio than in other regions. These spectral regions of advantageous signal-to-noise-ratio may vary dynamically as a function of time. Because many natural masking stimuli are characterized by this type of modulation, such as a competing speech background, it is of interest to determine how well observers utilize cues to speech identification from different spectral regions that occur asynchronously as opposed to synchronously. Recent work in automatic speech recognition has suggested that use of such distributed glimpses could be important in modeling the processing of speech in noise (Cooke, 2004).

From a purely informational perspective, it is plausible that asynchronous cues to speech identification might provide information absent in the case of coherent masker AM, where all such cues would occur synchronously. For example, speech sounds such as /t/ that are spectrally broad, and so in some sense redundant across frequency, might be detected efficiently based on samples from a small number of spectral channels, with additional samples from other channels providing little or no new information. In that case, a distribution of samples across frequency regions over time would provide new information about an ongoing speech stream, and clustering of samples would provide primarily redundant information.

There are some suggestions in the literature that the auditory system may be relatively efficient at piecing together or combining asynchronous cues to speech identification. For example, data from Arai and Greenberg (1998) indicated that the perception of speech that is filtered into multiple narrow spectral bands remains good for temporal misalignments across bands of up to 140 ms (see also Greenberg and Arai, 1998). Festen and Plomp (1990) showed that an AM speech-shaped noise was no less effective at masking sentence material than a speech-shaped noise with different patterns of AM applied to the portions below and above 1000 Hz. A recent study by Buss *et al.* (2003) is also relevant to this question. That study investigated the effect of AM coherence on masked speech recognition and included conditions in which speech was filtered into nine spectral bands that were

[a]Electronic mail: ebuss@med.unc.edu

then amplitude modulated either coherently or incoherently via multiplication with envelopes extracted from narrow bands of noise. Results indicated no masked recognition advantage for coherently modulated bands of speech relative to incoherently modulated bands. While this result is consistent with processing that combines cues to speech identification across time and frequency, one factor limiting interpretation of the incoherent AM results is the fact that the incoherent AM produced *un*correlated rather than *negatively* correlated patterns of modulation across frequency in this paradigm. Thus, the degree of across-frequency asynchrony of available speech cues was not under tight experimental control.

This factor was more rigorously controlled in a study by Howard-Jones and Rosen (1993), where the explicit intent was to study combination of asynchronous cues to consonant identification in a VCV context. In that paradigm, a pink-noise masker was filtered into contiguous bands and amplitude modulated via multiplication with a 50% duty cycle raised square wave, with neighboring bands modulated either in phase or 180° out of phase. Results in the antiphasic condition were compared with those obtained for either subset of coherently AM noise bands; greater sensitivity in the *out-of-phase* condition than for either subset was interpreted as reflecting combination of cues to speech identification. Evidence of combination of asynchronous information was observed only when masker bandwidths were relatively wide, that is, with a small number of bands. These results could have been influenced by several factors other than the ability to combine asynchronous cues to speech identification, however, such as upward spread of masking across juxtaposed bands of masking noise.

II. EXPERIMENT 1

Experiment 1 used a method very similar to that of Howard-Jones and Rosen (1993), but instead of manipulating the distribution of cues to consonant identification via masker AM, epochs of improved signal-to-noise ratio were manipulated via AM of the speech signal, at either 10 or 20 Hz. This method was chosen because it limits the degree to which upward spread of masking could obscure speech information in neighboring spectral bands. It is possible that portions of the speech signal corresponding to AM maxima would cause upward spread of excitation into spectral regions of signal corresponding to AM minima. However, this effect would not act to obscure information because such information would have already been eliminated via AM. This method may therefore be more sensitive than that of Howard-Jones and Rosen (1993) in assessing the ability of the auditory system to integrate asynchronous cues to speech identification across time and spectral regions.

A. Methods

1. Observers

A total of 13 adult observers participated in experiment 1, ages 19 to 30 years (mean of 22 years). All observers had thresholds that were better than or equal to 20 dB HL at octave frequencies from 250 to 8000 Hz (ANSI, 1996). The participation of adults as paid observers was approved by the Human Subjects Institutional Review Board of the University of North Carolina at Chapel Hill, School of Medicine. Observers 1–7 participated in the first set of conditions, utilizing the 10-Hz AM rate. Observers 3 and 4 then completed the second set of conditions, utilizing the 20-Hz AM rate, as did six additional observers. No attempt was made to recruit naive observers for the primary experiments, and several observers had participated in pilot versions of this experiment.

2. Stimulus generation

An American English-speaking female produced 10 tokens each of 12 vowel–consonant–vowel (VCV) utterances, of the form /*a*/ C /*a*/ (e.g., /*aga*/), with approximately equal stress on both syllables. Consonant recognition is a good indicator of speech processing abilities because of the high information content of consonants (Assmann and Summerfield, 2004). The particular consonants selected were the same as those used in Howard-Jones and Rosen (1993), that is: b, d, f, g, k, m, n, p, s, t, v, and z. From this recording, five tokens of each were selected based on uniformity of prosody, amplitude, and stress across all stimuli. These tokens were 528–664 ms in duration (mean 608 ms). Prior to the experiment, the 60 tokens (5 replications of 12 VCVs) were scaled to have equal total rms level, and average presentation level of the tokens prior to AM manipulation was 60 dB SPL.

Stimuli were filtered into 2, 4, 8, or 16 contiguous bands, with edge frequencies logarithmically spaced from 0.1 to 10 kHz. This was achieved using MATLAB; each token was submitted to an FFT, the magnitudes of points outside the passband were set to zero, and the result was transformed back into the time domain via IFFT. After filtering each of the 60 speech tokens, the resulting stimuli were saved to disk in two files, one file containing the odd-numbered bands and one containing the even-numbered bands. These files were retrieved during the experiment and amplitude modulated via multiplication with a 50% duty cycle raised square wave; duty cycle transitions were smoothed with 5-ms $\cos^2$ ramps. The first data set was collected using a 10-Hz modulation rate, the same rate used by Howard-Jones and Rosen (1993). A second data set was collected using a 20-Hz rate of modulation in order to assess the generality of results with respect to modulation rate.

The five modulation conditions are illustrated in Fig. 1, using the example of four bands. In the *just-odd* conditions, AM of the odd-numbered bands started with an ''on'' epoch (AM maximum), and the even-numbered bands were omitted entirely. In the *just-even* conditions, AM of the even-numbered bands started with an ''off'' epoch (AM minimum), and the odd-numbered bands were omitted entirely. In the *out-of-phase* conditions, the AM of the odd-numbered bands started with an ''on'' epoch and the even-numbered bands with an ''off'' epoch, such that these stimuli were the sum of comparable *just-odd* and *just-even* stimuli. For the two *comod* conditions, all bands were incorporated, both the odd- and even-numbered subsets, and phase of AM was coherent across these sets. In the *comod-odd* condition, AM began with an ''on'' epoch, such as that used in the *just-odd* condition. In the *comod-even* condition, AM began with an ''off'' epoch, such as that used in the *just-even*. Because the stimuli were filtered into contiguous bands, the *comod-odd*

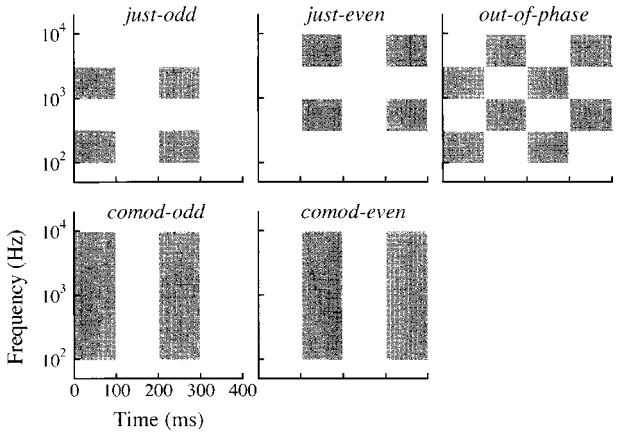


FIG. 1. Schematic of the four-band modulation conditions. Shaded areas indicate ''on'' periods.

and *comod-even* conditions generated with different numbers of bands would have resulted in identical stimuli; therefore, a single threshold estimate is reported for each comodulation condition. Figure 2 is a spectogram of a token of /*asa*/, as presented in the 10-Hz, four-band *out-of-phase* condition. This figure illustrates the effects of AM on the temporal and spectral boundaries of ''on'' epochs.

All speech signal bands were digitally summed prior to presentation via digital-to-analog converter at a rate of 44.1 kHz. The masker was a pink noise, following Howard-Jones and Rosen (1993), generated by routing the output of a Gaussian noise generator through a programmable digital filter; this masker has the advantage of providing equal masking for equal log bands. The level of the masker was adjusted via programmable attenuator, and played continuously over the duration of a threshold track. Signal and masker streams were summed and presented over the left transducer of a Sennheiser HD265 headphone. While some data indicate right-ear advantage for speech processed under challenging conditions (e.g., faster reaction time in the presence of a competing message; Treisman and Gerren, 1968), it is unlikely that the ear of presentation played a role in the results presented here.

3. Procedures

During the experiment, the observers heard a randomly selected VCV token. They were then visually presented with the 12 possible consonants and asked to enter a response via the keyboard. No explicit feedback was provided about the identity of the token so as not to encourage the use of cues other than those used under more natural listening conditions. The level of the masker was adjusted using a three-up, one-down adaptive rule, estimating the masker level necessary to obtain 79%-correct identification. The first level adjustment was 8 dB, the second 4 dB, and the remaining eight reversals were made in steps of 2 dB. The threshold estimate was the mean masker level at the last eight track reversals. Three such estimates were obtained in each of 14 conditions—the two *comod* conditions and the three noncomodulation conditions (*just-even*, *just-odd*, and *out-of-phase*) at four band frequencies—completed in random order.

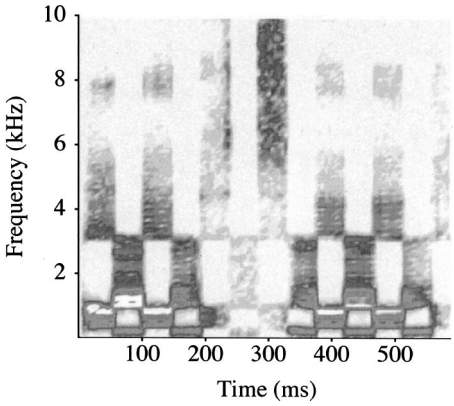


FIG. 2. Example spectogram of /*asa*/ in the four-band *out-of-phase* condition.

B. Results and discussion

Figure 3 shows the mean masker level at threshold in each of the 10-Hz AM conditions, as indicated along the abscissa. Results are plotted separately for each observer and the mean across observers, with symbols indicating the numbers of bands present. Symbols are connected by lines to highlight trends in the data, but it should be kept in mind that the ordering of conditions along the abscissa is arbitrary. Any track for which masker levels dipped consistently below masker audibility, defined as two or more occurrences in a single track, was interpreted as floor performance. Threshold of masker audibility was estimated based on the average of three detection threshold estimates obtained from each of four observers using a three-alternative forced-choice paradigm that incorporated the same procedures described above. The resulting estimate of masker audibility was 12.5 dB SPL. Any VCV identification threshold estimate falling below that level was set to 12.5 dB SPL for the purposes of plotting and data analysis.

The 10-Hz AM data were not consistently different within observer in the two *comod* conditions (paired, 2-tailed t-test, $p=0.60$). In this data set, there was a slight advantage associated with coherent AM, where cues are synchronous across frequency, as opposed to incoherent AM: thresholds in the pooled *out-of-phase* conditions were on average 4.0 dB lower (poorer) than those in the pooled *comod* conditions. A paired t-test comparing the mean *out-of-phase* and mean *comod* thresholds for each observer supports this observation ($p<0.005$). Masker level at threshold in the *out-of-phase* conditions was on average 20.5 dB higher than that in the *just-odd* condition and 11.2 dB higher than that in the *just-even* condition, indicating better sensitivity when both subsets of bands were present. This difference was confirmed with a repeated measures ANOVA, including three levels of AM pattern (*out-of-phase*, *just-odd*, and *just-even*) and four levels of band number (2, 4, 8, and 16). The analysis indicated significant main effects of AM pattern ($F_{2,12}=80.42, p<0.0001$) and band number ($F_{3,18}=6.40, p<0.005$), and a significant interaction ($F_{6,36}=5.72, p<0.0005$). A simple contrast between the *out-of-phase* and *just-odd* conditions revealed a significant difference ($F_{1.6}=112.81, p<0.0001$), as did the contrast between *out-of-phase* and *just-even* conditions ($F_{1.6}=53.27, p<0.0005$). These results suggest that the brief epochs of speech do not

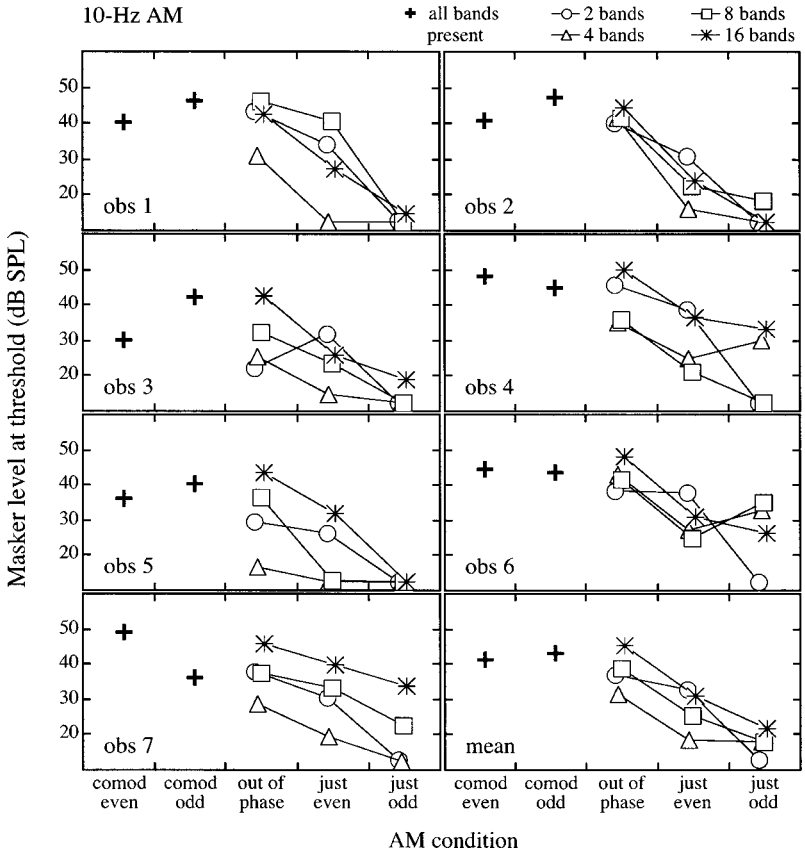


FIG. 3. The mean masker level at VCV identification threshold with a 10-Hz AM signal in the five AM conditions (abscissa) for each of the seven observers (panels), with the mean across observers shown in the bottom right panel. Symbols indicate thresholds for the different numbers of bands and for the *comod* results that are representative of all numbers of bands.

need to be presented synchronously across frequency in order for observers to combine and utilize cues in both subsets of bands. A test of simple main effects examined the effect of condition for each number of bands: in all but one case, results indicate better sensitivity in the *out-of-phase* condition than either the *just-even* or *just-odd* conditions ($p<0.05$, with Bonferroni correction for multiple comparisons). The exception was the two-band, *just-even* vs *out-of-phase* difference, which failed to reach significance ($p=0.188$ without adjustments for multiple comparison).

Results for the 20-Hz AM are shown in Fig. 4 using the same plotting conventions as those used in Fig. 3. These results are broadly similar to those obtained with the 10-Hz AM rate. As with the 10-Hz AM data, thresholds were not significantly different in the two *comod* conditions (paired 2-tailed t-test, $p=0.32$). Thresholds in the *out-of-phase* conditions were on average 1.1 dB higher (better) than those in the *comod* conditions, but this difference failed to reach significance (paired 2-tailed t-test, $p=0.50$). Masker level at threshold in the *out-of-phase* conditions was higher than those in either the *just-odd* or *just-even* conditions, by 23.9 and 19.3 dB, respectively. This was confirmed with a repeated measures ANOVA, including three levels of AM pattern (*out-of-phase, just-odd,* and *just-even*) and four levels of band number (2, 4, 8, and 16). This resulted in significant main effects of AM pattern ($F_{2,14}=140.01, p<0.0001$) and band number ($F_{3,21}=9.27, p<0.0005$), and a significant interaction ($F_{6,42}=4.13, p<0.005$). A simple contrast between the *out-of-phase* and *just-odd* conditions resulted in a significant difference ($F_{1,7}=231.54, p<0.0001$), as did the contrast between *out-of-phase* and *just-even* conditions ($F_{1,7}=188.18, p<0.0001$). A test of simple main effects examined the effect of condition for each number of bands: for all numbers of bands, results indicate better sensitivity in the *out-of-phase* condition than either the *just-even* or *just-odd* conditions ($p<0.05$, Bonferroni correction for multiple comparisons).

For both the 10- and 20-Hz rates, masker level at threshold for the *just-odd* and *just-even* conditions tended to rise with increasing numbers of bands, perhaps because an increasingly broad spacing of bands results in more efficient sampling of the full spectrum of the speech. This can be seen in Fig. 5, where average masker level at threshold is plotted as a function of number of bands for the 10-Hz and 20-Hz conditions. One exception to this trend is the relatively high masker level at threshold in the two-band, *just-even* condition. This finding is consistent with the fact that the Articulation Index for speech shows a ''center of gravity'' that falls in the region of 1.9 kHz (French and Steinberg, 1947). When the VCV tokens are filtered into two bands, the even-numbered band (band 2) spans 1–10 kHz, comprising more than half of the relevant speech information, possibly accounting for good performance in this condition. In terms of stimulus energy, however, the bandwidths used here partitioned the stimulus into approximately equivalent bands; for the two-band case, band 2 (1–10 kHz) was only 1.8 dB more intense on average than band 1 (0.1–1 kHz). Level calculations for the other numbers of bands support the assumption that signal level for the *just-odd* and *just-even* conditions was comparable, with average differences at around 1 dB. The pattern of *out-of-phase* thresholds as a function of number of bands shown in Fig. 5 is less clear, with some suggestion that

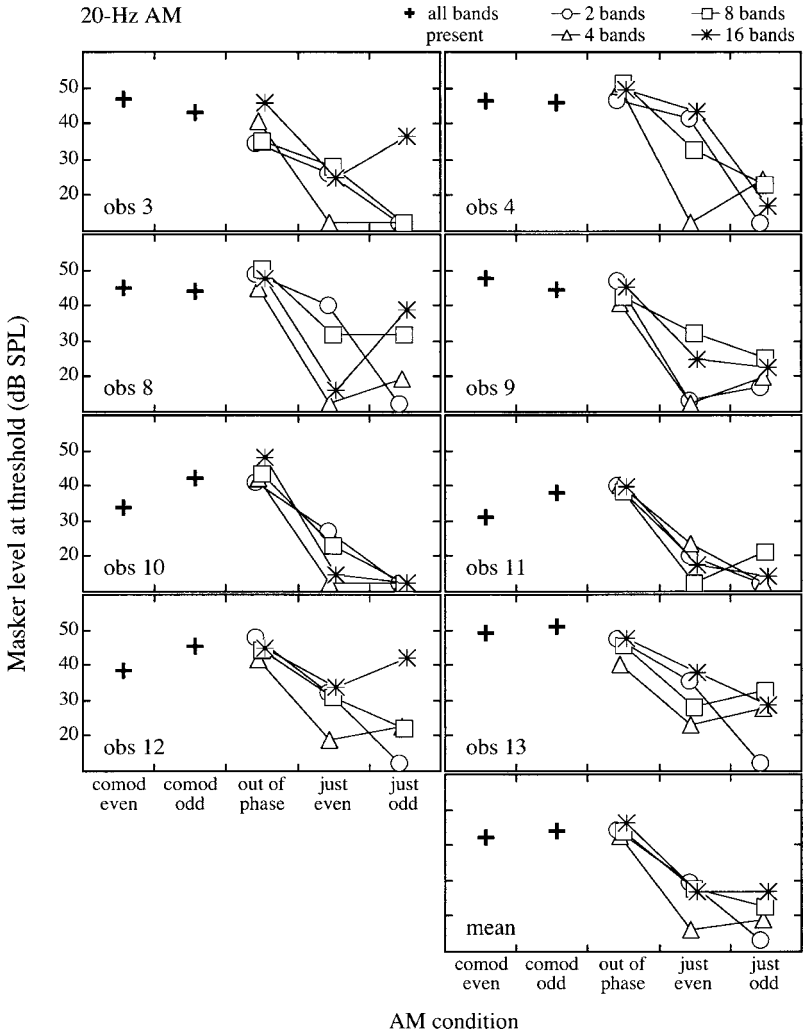


FIG. 4. The mean masker level at VCV identification threshold with a 20-Hz AM signal in the five AM conditions (abscissa) for each of the eight observers (panels), with the mean across observers shown in the bottom right panel. Symbols indicate conditions employing different numbers of bands, as in Fig. 3.

these thresholds follow the same pattern as those of the *just-even*, particularly for the 10-Hz rate: that is, masker level at threshold tends to rise as a function of number of bands, with the exception of the two-band case.

Additional supplemental data were collected to determine the relative contributions of the various subsets of bands. This testing used stimuli in the *just-even* and *just-odd* conditions, with the exception that the speech materials were not amplitude modulated. Five naive observers participated, all fitting inclusion criteria described above and all analogous demographically to the observers tested previously. Testing was carried out following the procedures described for the two experiments reported above. Results are reported in Table I. Thresholds for the even-numbered bands were approximately constant at 45.5 dB. Thresholds for the odd-numbered bands, however, varied substantially with number of bands, increasing monotonically from 22.3 dB for the two-band case to 47.7 for the 16-band case. These results are consistent with the speculation that the failure to find evidence of integration of asynchronous cues in the two-band case for 10-Hz AM was due to a very weak cue in the band associated with the *just-odd* two-band condition. One additional data point collected as supplemental data was the threshold for the full complement of bands present without AM. The average threshold in this condition was 54.7 dB. This value can be compared to thresholds for the comodulated conditions, averaging 42.0 and 43.0 dB for 10 and 20 Hz AM, respectively; this indicates that the AM applied to the VCV stimuli in these conditions introduced substantial susceptibility to masking.

III. EXPERIMENT 2

The results of experiment 1 are consistent with the conclusion that listeners are able to integrate cues to consonant identification across frequency over time for even the highest number of bands tested. This evidence is based on the finding of higher masker levels at threshold for the *out-of-phase* conditions compared to either the *just-odd* or the *just-even* conditions. An alternative explanation not ruled out by the procedures of experiment 1, however, is that subsets of bands, either even- or odd-numbered bands, are highly variable in terms of the cue quality for individual tokens, and that performance in the *out-of-phase* condition can be explained solely in terms of reliance on the better of the two subsets of cues. This possibility is bolstered by findings that cues to phoneme identification are not distributed uniformly across frequency, with different classes of consonants, for example, being associated with transient cues in different frequency bands (Ghitza, 1993). If some of the tokens in

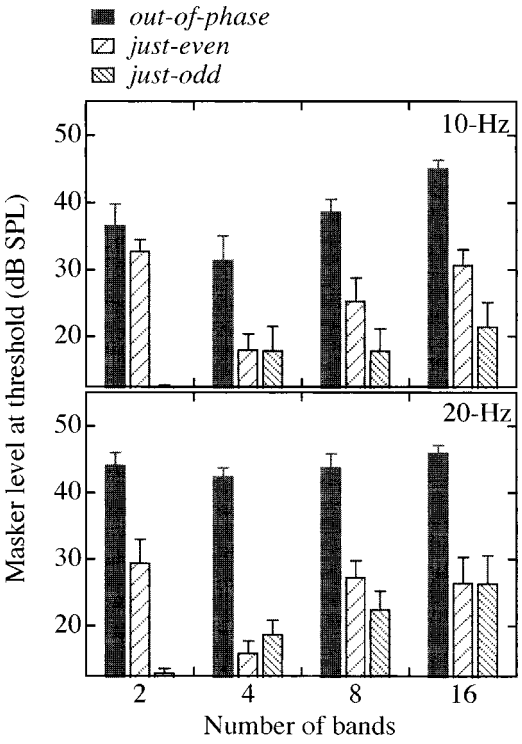


FIG. 5. Mean threshold across observers for each number of bands is plotted for the 10- and 20-Hz AM rates. Bars with the solid fill indicate the *out-of-phase* condition, rising hatches indicate the *just-even* condition, and falling hatches indicate the *just-odd* condition. Error bars show 1 standard error of the mean.

experiment 1 were associated with high-quality cues in the even-numbered bands and others in the odd-numbered bands, then results from a track where tokens were interleaved would give the impression of integration, when in fact performance was based on just the subset of bands providing the better cue.

This possibility was examined in experiment 2, where thresholds were obtained for individual tokens, so that *out-of-phase* thresholds could be compared to *just-even* and *just-odd* thresholds for the same tokens. A similar pattern of results as that obtained in experiment 1 would be expected if across-frequency combination of cues as a function of time was responsible for those results. Dominance of either subset of bands in the original data set would be evidenced by consistent patterns in those conditions across observer, as well as comparable performance in the *out-of-phase* and either the *just-odd* or *just-even*.

TABLE I. Average masker level at threshold for a subsidiary data set collected with stimuli that were not AM. Standard deviation across six listeners is noted below each mean.

	Even	Odd	Both subsets
2 bands	45.63	22.32	54.68
	(4.25)	*(5.99)*	*(1.79)*
4 bands	45.24	38.68	
	(6.66)	*(5.94)*	
8 bands	45.92	43.96	
	(4.22)	*(2.80)*	
16 bands	45.81	47.74	
	(2.94)	*(2.21)*	

A. Methods

1. Observers

Five observers participated in experiment 2, all meeting the inclusion criteria outlined above. Three observers (Obs 1, Obs 6, and Obs 11) had previously participated in experiment 1. The two remaining observers had previously participated in psychoacoustical testing, but had no prior experience with the speech stimuli used here.

2. Procedures

The procedures and stimuli were identical to those described above with the following exceptions. In this experiment a single token of each of the 12 VCV stimuli was randomly selected from the five replicate productions. Over the course of a threshold estimation track each of these 12 tokens was presented 30 times, with order randomly determined for each track. Each of the tokens was associated with a three-up, one-down track, such that the masker level for each trial was determined by the track associated with the token presented on that trial. The first level adjustment of each track was 8 dB, the second 4 dB, and all remaining reversals were made in steps of 2 dB. The masker level at the last even number of track reversals, excluding the first two, was averaged to produce a threshold estimate. Using this method, estimates were based upon an average of 4–6 reversals. In several instances, fewer than four reversals were obtained. This occurred because of repeated failure to correctly identify a token despite the fact that the masker fell below the audibility threshold. Following experiment 1, a threshold of 12.5 dB SPL was used for plotting and analysis in these cases. Three sets of threshold estimation tracks were obtained for each token in *out-of-phase*, *just-even*, and *just-odd* conditions, all employing four-band filtered stimuli and the 10-Hz AM rate.

B. Results and discussion

Figure 6 shows thresholds as a function of VCV token for a representative observer (Obs 6) and for the mean across the five observers. Stars over the data for "m" and "v" in the top panel highlight cases where the *out-of-phase* condition failed to produce greater resistance to masking than either subset of bands presented alone. Note the large variability across tokens in this observer's data. Other observers showed comparable variability, though the pattern of results was not consistent. For example, in the data of Obs 1 there were two cases for which the *out-of-phase* results failed to produce greater resistance to masking than either of the two subsets of bands presented alone; those tokens were "f" and "s". Across observers this failure occurred no more than twice for any single token.

The trend towards greater resistance to masking in the *out-of-phase* condition as compared to either subset of bands presented alone was confirmed with a repeated measures ANOVA. For each condition and for each observer the threshold representing the best performance in either the *just-even* or the *just-odd* conditions was selected, and those data were compared to thresholds from the *out-of-phase* condition. This resulted in significant main effects of condition

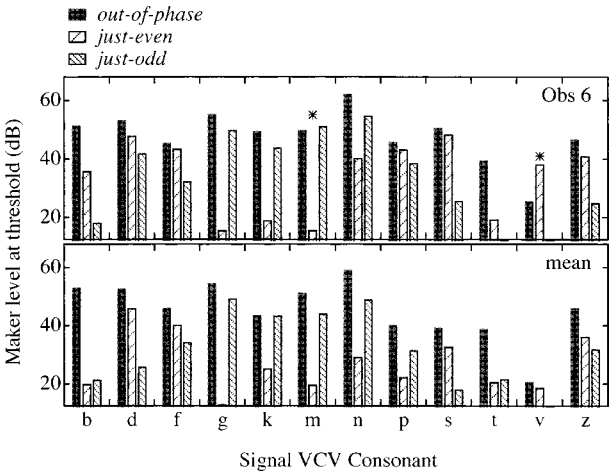


FIG. 6. The mean masker level at threshold for a representative observer (Obs 6) and mean across observers ($n=5$). As in Fig. 5, bars with the solid fill indicate the *out-of-phase* condition, rising hatches indicate the *just-even* condition, and falling hatches indicate the *just-odd* condition. Stars appear in the top panel above data for conditions in which resistance to masking in the *out-of-phase* condition fails to exceed that in the analogous *just-even* or *just-odd* conditions.

($F_{1.4}=13.45, p<0.05$) and token ($F_{11.44}=9.35, p<0.0001$), and a significant interaction ($F_{11.44}=2.60, p<0.05$). These results suggest that combination of asynchronously presented cues occurs for at least some of the VCV tokens in this experiment; the significant interaction, however, suggests that this effect is not uniform across consonants.

While the findings presented here rule out the possibility that performance in the *out-of-phase* condition of experiment 1 is due to cues present in the subset of bands comprising either the *just-odd* or *just-even* conditions, they do not rule out the possibility that observers were dynamically shifting attention between subsets of bands over the course of a single token. The second possibility is related to the idea that improvement in detection threshold for a pure tone as a function of duration is due to combination ''multiple looks'' at the output of a relatively brief integration window; this is in contrast to detection based on the output of a single, more prolonged integration window (Viemeister and Wakefield, 1991). A long integration window, with a time constant on the order of hundreds of milliseconds, would not be appropriate to model combination of information for the type of stimuli used here because the temporal and spectral fluctuation patterns of speech that allow tokens to be identified are preserved in the percept. In speech psychophysics ''integration'' is often taken to imply optimal combination of cues across channels prior to labeling of the speech sample (e.g., Braida, 1991), but because no attempt was made here to calculate optimal combination of cues, results must be interpreted with a less stringent definition of integration—as simply reflecting combination of cues.

IV. GENERAL DISCUSSION

The experiments described here were carried out to assess whether brief cues to the identity of a consonant presented in a VCV context could be combined across time as well as across frequency. One reason that this question is of interest is that such spectro-temporal processing may contribute to the ability of observers to understand speech in noisy environments. The results obtained here extend the work of Howard-Jones and Rosen (1993) by providing epochs of minimally masked speech information via modulation of the speech stimulus, whereas in the previous study the masker was amplitude modulated. In the AM speech conditions tested here, there is evidence of combination of information across time for even the highest number of bands, where performance in the *just-even* and *just-odd* conditions failed to reach that obtained in the *out-of-phase* condition. This finding is in contrast to Howard-Jones and Rosen's results, where the most convincing evidence for combination of asynchronous speech information was found in their two-band condition. It is interesting to note that the 10-Hz, two-band condition was the only condition in experiment 1 that did not show evidence of combination of asynchronous cues. Supplemental data collected for stimuli that had not received AM support the hypothesis that this result is due to unequal distribution of cues across frequency. That is, failure to demonstrate integration in the two-band case is likely due to the relatively poor cue in the band spanning 0.1 to 1 kHz.

Howard-Jones and Rosen report that even for the two-band condition, where evidence of combination of asynchronous cues was found, performance was considerably better in the presence of coherently than incoherently modulated noise: the signal-to-noise ratio at threshold was approximately 8 dB better in the comodulated noise. In contrast, differences between thresholds for the synchronous and asynchronous AM speech reported here were small (4 dB for the 10-Hz AM) or nonsignificant (1 dB for 20 Hz) in the present study. There are at least three reasons why the coherently AM noise may have yielded substantially better performance than the incoherently AM noise in the Howard-Jones and Rosen study. One possibility is that the reduced performance in the incoherent AM condition reflects a reduced ability to integrate asynchronous speech information. However, this explanation is at odds with the present findings of a more modest benefit of coherence, suggesting that the ability to integrate asynchronous information is not the sole factor contributing to the Howard-Jones and Rosen finding. A second possibility is that upward spread of masking associated with the incoherently AM noise limited performance in this condition. A third possibility is that the better performance in the coherent AM noise condition reflects a contribution of comodulation masking release (Hall *et al.*, 1984). Although comodulation masking release effects appear to be relatively small for suprathreshold signals, including speech (Grose and Hall, 1992; Hall *et al.*, 1997; Buss *et al.*, 2003; Kwon, 2002), some contribution of this phenomenon to the results of Howard-Jones and Rosen (1993) cannot be ruled out.

Regardless of how the discrepancy between AM speech and AM masker data is interpreted, these results raise the question of whether integration of asynchronous cues across a large number of bands has a material role in speech processing under natural listening conditions. In some ways the AM masker manipulation more closely resembles a natural stimulus for which asynchronous ''glimpses'' of speech would be available across disparate frequency regions, such

as a speech background noise, and yet the AM masker manipulation used by Howard-Jones and Rosen failed to provide evidence for extensive asynchronous integration. One factor to consider is the extreme nature of the stimuli used in that manipulation, with transitions between temporal and spectral regions of masking that are more abrupt than those typically found under natural listening conditions. A more naturalistic masker would be characterized by more gradual spectral and temporal transitions that would be less subject to nonsimultaneous masking or spread of masking. Whereas the conditions examined here also fall short of measuring performance under natural conditions, the present results nevertheless demonstrate that integration of asynchronous information can be quite effective across a range of conditions. More work is needed to establish whether similar results could be obtained with stimuli more closely resembling those observed in natural listening environments.

V. CONCLUSIONS

The results reported here demonstrate that the auditory system is able to combine asynchronous speech spectral information for a wide range of spectral band number (up to 16 bands) and for both 10- and 20-Hz rates of modulation. Whereas the results from several observers show comparable performance for the *comod* conditions and for the *out-of-phase* conditions, particularly for the 20-Hz AM, other observers showed better performance in coherent modulation conditions. With the exception of the two-band, 10-Hz AM conditions, sensitivity in the *out-of-phase* conditions exceeded that in either the *just-odd* or *just-even* conditions, consistent with the conclusion that brief epochs of speech information are being integrated across spectral regions over time in the *out-of-phase* conditions. More extensive data collected for a subset of VCV tokens support this conclusion. Overall, the results of this investigation are consistent with an interpretation that observers are able to utilize cues from different spectral regions that occur asynchronously as opposed to synchronously across frequency. It is possible that this ability contributes to speech recognition in noisy environments.

ACKNOWLEDGMENTS

This work was supported by NIH NIDCD RO1 DC00418. A subset of these data was presented at the summer 2002 meeting of the Acoustical Society of America in Chicago [Hall, Buss, and Grose (2002)]. Discussions with Peter Gordon were integral to developing the approach presented here. We thank Ken Grant, Stuart Rosen, and Steve Greenberg for their helpful comments.

ANSI (**1996**). ANSI S3-1996, ''American National Standards Specification for Audiometers'' (American National Standards Institute, New York).

Arai, T., and Greenberg, S. (**1998**). ''Speech intelligibility in the presence of cross-channel spectral asynchrony,'' in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Seattle, pp. 933–936.

Assmann, P. F., and Summerfield, Q. (**2004**). ''The perception of speech under adverse acoustic conditions,'' in *Speech Processing in the Auditory System*, edited by S. Greenberg, W. A. Ainsworth, A. N. Popper, and R. R. Fay, Springer Handbook of Auditory Research Vol 18 (Springer, Berlin).

Braida, L. D. (**1991**). ''Crossmodal integration in the identification of consonant segments,'' Q. J. Exp. Psychol. A **43A**, 647–677.

Buss, E., Hall, III, J. W., and Grose, J. H. (**2003**). ''Effect of amplitude modulation coherence for masked speech signals filtered into narrow bands,'' J. Acoust. Soc. Am. **113**, 462–467.

Carhart, R., Tillman, T., and Johnson, K. (**1966**). ''Binaural masking of speech by periodically modulated noise,'' J. Acoust. Soc. Am. **39**, 1037–1050.

Cooke, M. (**2004**). ''Glimpsing speech,'' J. Phonetics (in press).

Dirks, D. D., Wilson, R. H., and Bower, D. R. (**1969**). ''Effect of pulsed masking on selected speech materials,'' J. Acoust. Soc. Am. **46**, 898–906.

Festen, J. M., and Plomp, R. (**1990**). ''Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing,'' J. Acoust. Soc. Am. **88**, 1725–1736.

French, N. R., and Steinberg, J. C. (**1947**). ''Factors governing the intelligibility of speech sounds,'' J. Acoust. Soc. Am. **19**, 90–118.

Ghitza, O. (**1993**). ''Processing of spoken CVCs in the auditory periphery. I. Psychophysics,'' J. Acoust. Soc. Am. **94**, 2507–2516.

Greenberg, S., and Arai, T. (**1998**). ''Speech intelligibility is highly tolerant of cross-channel spectral asynchrony,'' J. Acoust. Soc. Am. **103**, 3057.

Grose, J. H., and Hall, III, J. W. (**1992**). ''Comodulation masking release for speech stimuli,'' J. Acoust. Soc. Am. **91**, 1042–1050.

Gustafsson, H. A., and Arlinger, S. D. (**1994**). ''Masking of speech by amplitude-modulated noise,'' J. Acoust. Soc. Am. **95**, 518–529.

Hall, III, J. W., Buss, E., and Grose, J. H. (**2002**). ''Spectral integration of synchronous and asynchronous cues to consonant identification,'' J. Acoust. Soc. Am. **111**, 2428.

Hall, III, J. W., Grose, J. H., and Dev, M. B. (**1997**). ''Signal detection and pitch ranking in conditions of masking release,'' J. Acoust. Soc. Am. **102**, 1746–1754.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). ''Detection in noise by spectro-temporal pattern analysis,'' J. Acoust. Soc. Am. **76**, 50–56.

Howard-Jones, P. A., and Rosen, S. (**1993**). ''Uncomodulated glimpsing in 'checkerboard' noise,'' J. Acoust. Soc. Am. **93**, 2915–2922.

Kwon, B. J. (**2002**). ''Comodulation masking release in consonant recognition,'' J. Acoust. Soc. Am. **112**, 634–641.

Miller, G. A., and Licklider, J. C. R. (**1950**). ''The intelligibility of interrupted speech,'' J. Acoust. Soc. Am. **22**, 167–173.

Treisman, A., and Geffen, G. (**1968**). ''Selective attention and cerebral dominance in perceiving and responding to speech messages,'' Q. J. Exp. Psychol. **20**, 139–150.

Viemeister, N. F., and Wakefield, G. H. (**1991**). ''Temporal integration and multiple looks,'' J. Acoust. Soc. Am. **90**, 858–865.

Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners

Peggy B. Nelson[a)]
Department of Communication Disorders, University of Minnesota, Minneapolis, Minnesota 55455

Su-Hyun Jin
Department of Communication Disorders, Department of Otolaryngology, University of Minnesota, Minneapolis, Minnesota 55455

(Received 5 February 2003; revised 12 February 2004; accepted 16 February 2004)

Previous work [Nelson, Jin, Carney, and Nelson (2003), J. Acoust. Soc. Am **113**, 961–968] suggested that cochlear implant users do not benefit from masking release when listening in modulated noise. The previous findings indicated that implant users experience little to no release from masking when identifying sentences in speech-shaped noise, regardless of the modulation frequency applied to the noise. The lack of masking release occurred for all implant subjects who were using three different devices and speech processing strategies. In the present study, possible causes of this reduced masking release in implant listeners were investigated. Normal-hearing listeners, implant users, and normal-hearing listeners presented with a four-band simulation of a cochlear implant were tested for their understanding of sentences in gated noise (1–32 Hz gate frequencies) when the duty cycle of the noise was varied from 25% to 75%. No systematic effect of noise duty cycle on implant and simulation listeners' performance was noted, indicating that the masking caused by gated noise is not only energetic masking. Masking release significantly increased when the number of spectral channels was increased from 4 to 12 for simulation listeners, suggesting that spectral resolution is important for masking release. Listeners were also tested for their understanding of gated sentences (sentences in quiet interrupted by periods of silence ranging from 1 to 32 Hz as a measure of auditory fusion, or the ability to integrate speech across temporal gaps. Implant and simulation listeners had significant difficulty understanding gated sentences at every gate frequency. When the number of spectral channels was increased for simulation listeners, their ability to understand gated sentences improved significantly. Findings suggest that implant listeners' difficulty understanding speech in modulated conditions is related to at least two (possibly related) factors: degraded spectral information and limitations in auditory fusion across temporal gaps. [DOI: 10.1121/1.1703538]

PACS numbers: 43.71.Ky, 43.66.Ts, 43.72.Dv, 43.66.Dc [PFA] Pages: 2286–2294

I. INTRODUCTION

Cochlear implant users commonly report difficulty understanding speech in noisy situations. It may be that device limitations, weak temporal pitch cues, the nature of pulsatile cochlear implant processing itself, or the lack of good spectral representation of speech results in difficulty segregating a target signal from a background noise. Nelson, Jin, Carney, and Nelson (2003) reported that listeners with cochlear implants did not demonstrate masking release from modulated maskers. Normal-hearing listeners obtain between 5 and 20 dB improvements in masked speech recognition when listening in gated noise compared to steady-state maskers of the same intensity (e.g., Bacon *et al.*, 1998). Implant listeners, in contrast, showed no difference in masked speech recognition performance when maskers were steady or gated noise, indicating that they did not take advantage of temporal gaps in gated noise. In addition, normal-hearing listeners who listened to four-band simulations of implant processing (after Shannon, Zeng, Kamath, Wygonski, and Ekelid, 1995, which greatly reduced speech spectral information) also failed to demonstrate masking release. These findings were in sharp contrast to well-established results from normal-hearing listeners (e.g., Gustafsson and Arlinger, 1994), who show significant masking release for gated maskers. Results from implant listeners were significantly poorer than those reported for listeners with cochlear hearing loss as well (e.g., Bacon, Opie, and Montoya, 1998). Nelson *et al.* (2003) reported that listeners with implants were significantly affected by background noise, even at favorable signal-to-noise ratios (SNRs), and did not demonstrate a benefit from short temporal "glimpses" of the quiet signal at any gate frequency, even when those temporal gaps in the noise were as long as 250 ms. Specific implant device-related processing characteristics (e.g., CIS versus SPEAK processing) did not seem to explain this finding, as listeners with three different processor types all showed the same reduced masking release. The authors proposed several possible explanations for the findings, including modulation masking, reduced auditory stream segregation or fusion abilities, and limited spectral resolution abilities of implant and simulation listeners. Some evidence in the literature suggests that implant listeners may show

[a)]Send correspondence to Peggy Nelson, Ph.D., 115 Shevlin Hall, Minneapolis, Minnesota 55455; Phone: (612) 625-4569; Fax: (612) 624-7586; E-mail: nelso477@umn.edu

significant modulation masking when speech signals are presented in background noise. Others have proposed that implant listeners have a difficult time segregating signals because of their reduced spectral resolution.

Kwon and Turner (2001) showed that normal listeners identifying consonant-vowel (CV) stimuli processed by implant simulations display increased susceptibility to high-frequency modulation masking when CV syllables are presented in gated noise. They hypothesized that cochlear implant users may be negatively affected by modulation masking because listeners with reduced spectral resolution have increased perceptual weighting for natural amplitude modulations in speech recognition (e.g., Hedrick and Jesteadt, 1996; Hedrick and Carney, 1997). In these studies, the amplitude cues were more important for hard-of-hearing (Hedrick and Jesteadt, 1996) and cochlear implant (Hedrick and Carney, 1997) listeners than for normal-hearing listeners. Thus, the modulated noise may directly interfere with the listener's perception of amplitude envelope cues at both the syllabic and segmental levels. Kwon and Turner evaluated the effects of modulated noise on listeners' understanding of spectrally impoverished signals using 12-band implant simulations. They found that under certain conditions, listeners obtained some masking release. However, a modulated high-frequency masker sometimes reduced consonant identification when compared to identification with an unmodulated masker. They concluded that high-frequency modulated maskers can cause interference in consonant recognition that may offset any benefit provided by the masking release. Apparently in the presence of high-frequency modulated noise, modulation masking caused consonant recognition errors by simulation listeners, and this effect may hold true for cochlear implant users.

In addition to their proposed susceptibility to modulation masking, implant listeners may demonstrate reduced ability to segregate auditory streams. Chatterjee and Galvin (2002) have shown that implant listeners can perform simple auditory stream segregation tasks. Implant listeners in their study were able to segregate rhythmic sequences presented on two electrodes. Greater electrode separation resulted in improved stream segregation, but listeners were able to stream sequences presented to the same electrode pair when those sequences had differing temporal envelopes. Results suggested that implant listeners can segregate streams of information when they differ in place (frequency) and/or modulation rate.

Grimault, Bacon, and Micheyl (2002) showed that normal-hearing listeners can segregate auditory streams on the basis of temporal fluctuations alone, in the absence of other spectral or temporal cues. Listeners generally segregated sequences into two streams when the difference in amplitude modulation (AM) rate between their noise carriers was greater than an octave. These results suggest that listeners (including cochlear implant users) may be able to successfully segregate sound streams based on the AM characteristics of each stream, regardless of the spectral content of the two sound stream carriers, as long as the modulation frequencies of the two sounds streams are at least an octave apart.

Qin and Oxenham (2003) recently reported that implant simulation listeners have significantly less making release when understanding speech in modulated versus steady noise. They compared normal-hearing listeners' speech recognition in steady speech-shaped noise, modulated speech-shaped noise, and single-talker speech interference under conditions of 4-, 8-, and 24-band implant simulations (after Shannon *et al.*, 1995). They found that increasing the number of spectral channels significantly improved simulation listeners' performance in modulated noise, but that even for simulation signals with 24 bands of spectral information, listener performance was significantly poorer than for the original speech signals. They hypothesized that spectral information is critical for segregation of speech from modulated noise, and that low-frequency pitch cues (that are reduced in simulation speech signals) are important for speech segregation.

In the current set of experiments, the various factors that may be related to reduced masking release in implant users were investigated, including informational masking, spectral resolution, and auditory fusion/segregation. Informational masking has been defined as threshold elevation due to non-energetic factors such as signal uncertainty, masker–stimulus similarity, or distraction from extraneous sounds (informational maskers) that do not physically mask the target signals (see Wightman *et al.*, 2003 and Durlach *et al.*, 2003). If implant users perceive the modulated maskers and speech signals as similar, the maskers may provide significant informational, rather than energetic, masking. The role of spectral resolution in masking release was proposed by our previous study and was investigated further by Qin and Oxenham (2003). Both studies hypothesize that reduced spectral information is related to reduced masking release from modulated noise. Auditory fusion and segregation of sentences from noise have also been proposed as important factors in understanding masking release. Qin and Oxenham also theorized that a strong pitch cue is necessary for normal segregation of speech from noise, and that implant listeners may not perceive the needed strong pitch cue.

These hypotheses are tested further with groups of normal and implant listeners. Listeners in the current experiments included cochlear implant users as well as listeners with normal hearing sensitivity who were tested for their understanding of speech that has been modified to reduce spectral cues (after Shannon *et al.*, 1995). We first investigated the roles of informational versus energetic masking in gated noise tasks (Experiment 1). In Experiment 1a, listeners were tested for their understanding of speech in gated noise (as in Nelson *et al.*, 2003) as the duty cycle of the noise was varied from 25% to 50% to 75%. Punch (1978) demonstrated that normal-hearing listeners show a systematic increase in masking as masker duty cycle increases from 25% to 50% to 75%, corresponding to the physical increase in the duration of the masker. Nelson *et al.* (2003) hypothesized that the spectrally impoverished signals experienced by implant and simulation listeners caused listeners to rely on amplitude envelope cues that were easily disrupted by gated noise. The disruption from gated noise might be due to the energetic masking of the gated masker (as observed in Punch), but the

disruption may be exaggerated in implant and simulation listeners who use the amplitude envelope cues for understanding. If listeners' performance in gated noise is directly related to the proportion of time that the gated noise is ''on,'' then performance is likely determined by the physical masking of the speech by the noise (energetic masking). If in the current study, however, the duty cycle of the noise has little effect on the amount of masking release, then it is likely that the effect of the gated masker was to disrupt the fusion of the auditory stream of information (causing ''informational masking''), rather than to physically mask the energy of the speech. Results of the investigation of masker duty cycle may shed some light on the role of informational masking in cochlear implant users.

An additional experiment using gated noise (Experiment 1b) tested the hypothesis that greater spectral resolution leads to increased masking release. Our previous work (Nelson *et al.*, 2003) showed that normal listeners obtain little to no masking release when they are listening to speech that has reduced spectral information. The relationship between the amount of spectral information and the amount of masking release is still unknown, however. If listeners with greater numbers of spectral channels are more successful in understanding speech in gated noise, it may be presumed that masking release in gated noise is directly related to the amount of signal spectral information available to the listeners. Qin and Oxenham (2003) recently demonstrated that increasing the number of spectral channels of implant simulations improves, but does not fully restore, masking release. In the current study this is explored further. Listeners with normal hearing sensitivity were presented with original full-spectrum speech stimuli, or with 12- or 4-band simulations of implant processing.

Experiment 2 focuses on the ability of implant and simulation listeners to understand interrupted sentences. When sentences are gated or interrupted by periods of silence, instead of noise, listeners who understand key words in sentences in the presence of the silent interruptions are presumed to have successfully grouped or fused the auditory information into a continuous speech stream. Research on understanding interrupted sentences was done by Miller and Licklider (1950) and by Bashford and Warren (1987), who showed that words and sentences interrupted by silent gaps were less intelligible than continuous speech. It was presumed that this task reflects auditory grouping or fusion (also called ''induction'' by Warren, 1999.) In Experiment 2, listeners were tested for their ability to understand sentences (in quiet) that had themselves been gated, or interrupted by silence. Listeners in Experiment 2a included implant users, normal-hearing listeners, and simulation listeners.

If listeners are more successful at fusing interrupted sentences with increasing spectral information, we presume that auditory fusion by implant listeners can be improved by increasing the amount of spectral information, or the number of spectral channels, available to the implant simulation listeners. In Experiment 2b, normal-hearing listeners were tested for their understanding of gated sentences using implant simulations with either 4 or 12 spectral bands.

II. METHODS

A. Stimuli

Speech stimuli consisted of IEEE sentence materials (IEEE, 1969) spoken by five male and five female talkers. Stimuli were recorded and digitized using a 20 kHz sampling rate. In addition, noise-vocoded simulations of the IEEE sentences were generated (after Shannon *et al.*, 1995) that simulate the processing strategies of cochlear implants. For the initial four-band implant simulations in Experiment 1a, speech stimuli from two talkers were filtered into four frequency bands (100–300, 300–500, 500–1700, and 1700–6000 Hz). The amplitude envelope was extracted from each band using the envelope extraction algorithm provided in Cool Edit Pro™. That envelope was then used to modulate the amplitude of narrowband noises having the same bandwidths. The resulting amplitude-modulated noise bands were recombined with their original relative amplitudes.

For Experiment 1, noise maskers were generated on the fly using the Tucker–Davis waveform generator (TDT WG1). The noise was passed through a Rane 30-band equalizer so that the spectrum of the resulting noise matched the long-term spectrum of the IEEE sentences. Noise maskers were presented either continuously (steady), or gated using 2 ms $\cos^2$-ramped gating that was also implemented on the fly.[1] Gate frequencies ranged from 1 to 32 Hz, while the duty cycle of the noise was systematically altered. Three duty cycle conditions were tested: 75% (the noise was on 75% of each cycle, followed by 25% silence), 50% (noise and silence each comprised 50% of each cycle), and 25% (the noise was on for only 25% of each cycle, followed by a 75% silent interval). The duration of the resulting noise bursts ranged from 8 ms (32 Hz gate frequency, 25% duty cycle) to 750 ms (1 Hz gate frequency, 75% duty cycle). Whereas the 75% condition provided the most physical masking of the stimuli, all three duty cycles produced noise bursts that presumably disrupt the envelope of the speech signal. Signal-to-noise ratios (SNRs) were +16, +8, 0, −8, or −16 dB, depending upon the listener and condition. Stimuli for Experiment 1b were either full-spectrum natural speech or speech processed through implant simulators (after Shannon *et al.*, 1995) that consisted of 4 to 12 bands. Filter settings are given in Table I. All stimuli for Experiment 1b were presented at −4 dB SNR.

For Experiment 2, sentences in quiet were gated with a 50% duty cycle using 4 ms $\cos^2$-ramped square gating. No noise was present for this experiment. Gate frequencies ranged from 1–32 Hz, resulting in regular bursts of speech that ranged in duration from approximately 14 (32 Hz gate frequency) to 500 ms (1 Hz gate frequency). Stimuli for Experiment 2b were either full-spectrum natural speech or speech processed through implant simulators with 4 and 12 bands, as shown in Table I.

B. Procedures

For all experiments, listeners were seated in the center of a sound-treated chamber. Speech signals were delivered diotically through two Bose 301 speakers (placed at ±45° azimuth) at an overall level of 65 dBA. Speech stimuli were

TABLE I. Filter cutoffs for 4- and 12-band simulations used in Experiments 1b and 2b.

4-band simulations:		
Band	Low-pass cutoff (Hz)	High-pass cutoff (Hz)
1	327	648
2	677	1341
3	1400	2773
4	2898	5739
12-band simulations:		
Band	Low-pass cutoff (Hz)	High-pass cutoff (Hz)
1	192	356
2	356	550
3	550	775
4	775	1036
5	1037	1342
6	1342	1700
7	1700	2116
8	2116	2602
9	2602	3169
10	3169	3830
11	3830	4600
12	4600	5500

presented in blocks of ten sentences, using all ten talkers in random order for each list. Sentences contained an average of five key words per sentence. All conditions were randomized prior to the beginning of each subject's testing. The listeners responded verbally to each sentence, and the experimenter scored the key words correct for each sentence, circling the correct answers on an answer form. Each listener's results (percent correct keywords) for each condition were later entered into computer database files.

For Experiment 1, keyword identification was evaluated in steady and gated noise. On each trial, the masking noise started first. The sentence began after a random delay that ranged from 10 to 100 ms. The noise was either steady or modulated, as described above. The level of the noise varied depending upon the condition being tested. The listeners in the implant and simulation groups heard the noise at +8- and +16-dB SNR. (Pilot testing with three very successful implant listeners indicated that performance was near 0% for all gate conditions at 0-dB SNR and lower.) Listeners in the normal group heard the noise at 0-, −8-, and −16-dB SNR. All listeners also completed two blocks of sentences in quiet.

For Experiment 2, key-word identification was evaluated for the gated sentences without noise. The sentences were presented either continuously, or were gated on and off, as described above.

C. Subjects

Subjects for the simulation experiments included two groups of eight young adult listeners with normal hearing sensitivity who were not familiar with the stimuli or the implant simulation algorithm. None of the normal-hearing listeners had participated in the previous experiments from Nelson *et al.* (2003). All were between the ages of 19 and 32 years and each participated in only one experiment.

Nine postlingually deafened users of cochlear implants were also tested. These were the same nine listeners that participated in the previous experiments (Nelson *et al.*, 2003). A detailed description of the implant users is shown in Table II. Their mean age was 49 years (range: 34 to 64 years), and their average length of deafness prior to implantation was 16 years (range 1 to 44 years). All listeners obtained significant open-set speech recognition from their implants and had worn their implants for more than 2 years (mean: 5 years, range 2 to 11 years). Three listeners used the Nucleus 22 device with the SPEAK processor, three used the Clarion 1.2 device with CIS processing, and three used the Clarion HiFocus device with the CIS processing strategy. Listeners in the implant group used their regular speech processors set to typical sensitivity with no noise reduction. At the beginning of each session, the users set the sensitivity while listening to practice lists, and they were instructed not to change the sensitivity setting.

TABLE II. Summary of implant subject characteristics.

Listener	CI/Processor	Age at test	Age at onset of deafness	Age at implantation
N12	Nucleus 22/SPEAK	53	32	42
N14	Nucleus 22/SPEAK	58	49	50
N32	Nucleus 22/SPEAK	34	5	29
C02	Clarion 1.2/CIS	42	18	37
C03	Clarion 1.2/CIS	53	22	49
C05	Clarion 1.2/CIS	47	42	43
C14	Clarion HiFocus/CIS	64	16	60
C15	Clarion HiFocus/CIS	42	33	40
C16	Clarion HiFocus/CIS	48	29	43

III. RESULTS AND DISCUSSION

A. Experiment 1a: Understanding speech in gated noise with varying duty cycles

Figure 1 shows the normal-hearing listeners' understanding of sentence keywords in gated noise at duty cycles of 25%, 50%, and 75%. Performance is plotted in percent

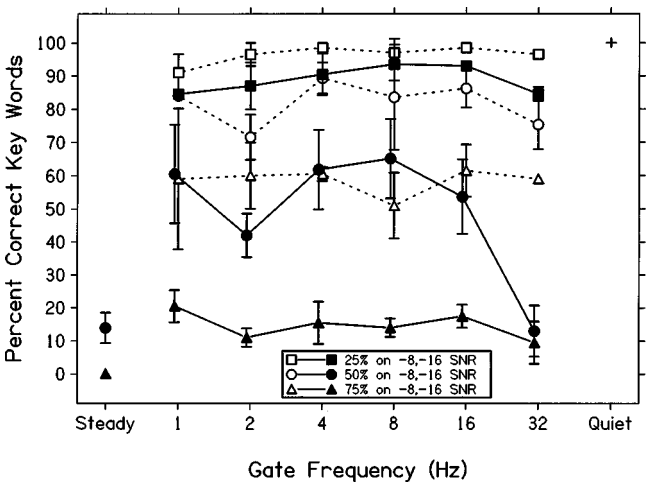


FIG. 1. Normal-hearing listeners' understanding of sentence key words in gated noise. Mean performance (±1 s.d.) is plotted in percent keywords correct as a function of gate frequency for noise presented at −16 dB (filled symbols) and −8-dB SNR. Gate frequencies range from 1 to 32 Hz. Performance in quiet is shown with a "+" symbol. Performance in steady noise is shown with a filled triangle at −16-dB SNR and with a filled circle at −8-dB SNR. A systematic effect of increasing duty cycle is seen.

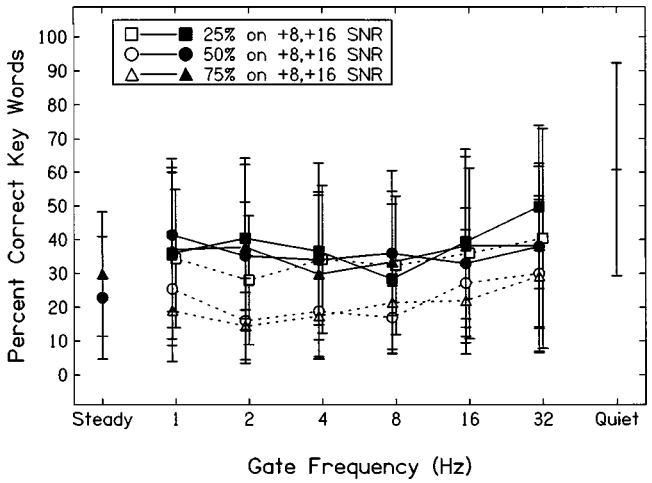


FIG. 2. Four-band simulation listeners' understanding of sentence keywords in gated noise. Mean performance (±1 s.d.) is plotted in percent keywords correct as a function of gate frequency for noise presented at +16 dB (filled symbols) and +8-dB SNR. Gate frequencies range from 1 to 32 Hz. Performance in quiet is shown with a "+" symbol. Performance in steady noise at +16-dB SNR is shown with a filled hourglass; performance in steady noise at +8-dB SNR is shown with an open hourglass. No systematic effect of masker duty cycle is noted.

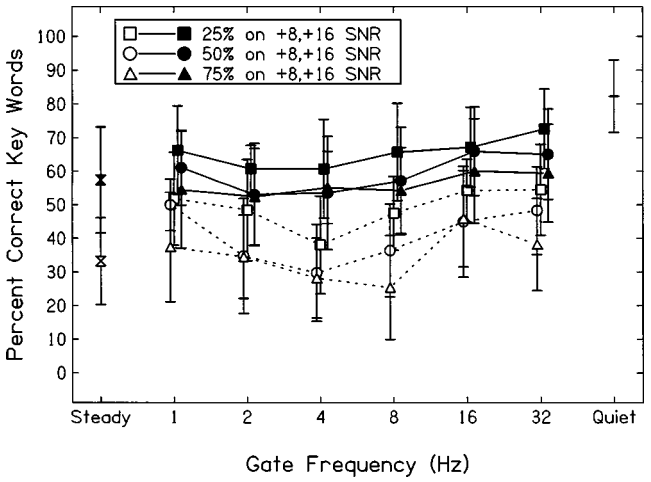


FIG. 3. Implant listeners' understanding of sentence keywords in gated noise. Mean performance (±1 s.d.) is plotted in percent keywords correct as a function of gate frequency for noise presented at +16 dB (filled symbols) and +8-dB SNR. Gate frequencies range from 1 to 32 Hz. Performance in quiet is shown with a "+" symbol. Performance in steady noise at +16-dB SNR is shown with a filled hourglass; performance in steady noise at +8-dB SNR is shown with an open hourglass. No systematic effect of masker duty cycle is noted.

key words correct as a function of gate frequency for noise presented at −16 dB (filled symbols) and −8-dB SNR (open symbols). Gate frequencies range from 1 to 32 Hz. For both SNRs, a predictable relationship between duty cycle and percent correct can be seen, demonstrating the energetic masking that occurs as the noise duty cycle increases from 25% onto 75% similar to the results from Punch (1978).

Normal-hearing listeners identify approximately 10% of keywords at 4 and 8 Hz gate frequencies at −16-dB SNR when the noise duty cycle is 75%. When the duty cycle is changed to 50%, they identify approximately 60% correct at the same gate frequencies. For the 25% duty cycle condition, they identify nearly 90% of the keywords correctly. The same trend is seen for the +8-dB SNR conditions, but the observed difference between 50% and 25% duty cycles is restricted presumably because performance approaches 100%. A repeated measure analysis of variance (ANOVA) indicated a significant difference in performance between each duty cycle for both SNRs ($F[2,10]=168.8, p<0.0001$). All Bonferroni corrected paired t tests were significant (at $p<0.01$). There is a clear systematic relationship between the energy of the masker and listeners' performance, suggesting that the energetic masking has a significant effect on performance in gated noise. Normal-hearing listeners' performance was best in noise with 25% duty cycle, moderate in noise with 50% duty cycle, and poorest in noise with 75% duty cycle. At the fastest modulation rates for −16-dB SNR, performance in the 50% duty cycle noise decreased compared to the slower rates. We presume this is due to the effects of forward masking, as noted in Nelson *et al.* (2003). At these fast rates and the highest noise levels, the gated noise may perceptually fill the silent interval, causing a sharp reduction in recognition.

In contrast, Figs. 2 and 3 show simulation and implant listeners' understanding of sentence keywords in gated noise, respectively. Performance is plotted in percent keywords correct as a function of gate frequency for noise presented at +16-dB (filled symbols) and +8-dB SNR (open symbols). Gate frequencies again range from 1 to 32 Hz. There is no apparent systematic effect of the noise duty cycle for implant or simulation listeners for either +8- or +16-dB SNR. Performance for the 25% duty cycle condition is slightly better than performance at 50% or 75% for the simulation listeners at +8-dB SNR ($F[2,22]=34.3, p<0.01$). Bonferroni-corrected paired t tests ($t[11]=0.04, p>0.05$) indicated that there is no significant difference in performance for 50% and 75% duty cycle noise. The implant listeners' performance is also slightly better for the 25% than for the 50% and 75% duty cycles. Bonferroni-corrected paired t tests ($t[17]=2.1, p>0.05$) indicated that there is no significant difference in performance for 50% and 75% duty cycle noise. Performance only differs by approximately 10% when the noise duty cycle varies from 25% "on" to 75% "on."

The data are summarized more clearly by collapsing across gate frequencies from 4 to 16 Hz, as shown in Figs. 4(a) (normal-hearing listeners) and 4(b) (simulation and implant listeners). Because there was no significant effect of gate frequency ($F[2,16]=3.78, p>0.05$), the data could be collapsed across these gate frequencies to more clearly illustrate the effect of the noise duty cycle.

Figure 4(a) (top panel) shows the mean performance of the normal-hearing listeners as a function of duty cycle (25% on, 50% on, 75% on, and steady) at two SNRs. Error bars represent one standard deviation. The clear trend for reduced performance with increasing duty cycle can be seen. Simply, the greater the amount of time the noise is on, the more the masking that occurs, suggesting that the effect of the gated noise is primarily energetic masking.

Figure 4(b) (lower panel) shows the mean performance of implant (black bars) and simulation (gray bars) listeners as a function of duty cycle (25% on, 50% on, 75% on, and steady) at two SNRs. Error bars represent one standard de-

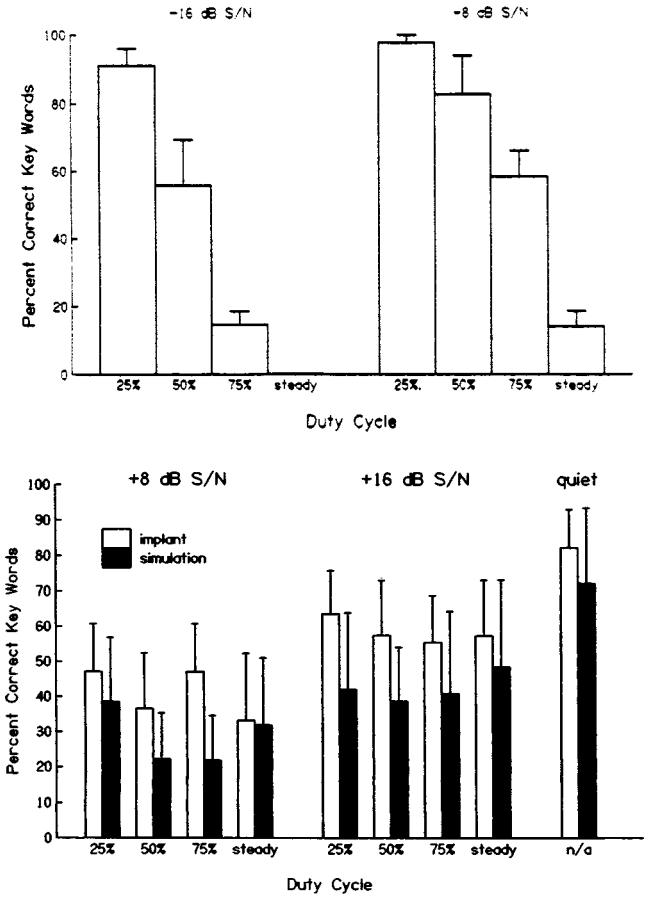


FIG. 4. In Panel (a) (top), mean performance of normal-hearing listeners is shown as percent keywords correct as a function of duty cycle (25% on, 50% on, 75% on, and steady) at two SNRs. Error bars represent one standard deviation. A clear trend for increased masking with increasing duty cycle can be seen. In (b) (bottom), mean performance of four-band simulation (black bars) and implant (gray bars) listeners is shown as percent keywords correct as a function of duty cycle (25% on, 50% on, 75% on, and steady) at two SNRs. Error bars represent one standard deviation. No clear trend for increased masking with increasing duty cycle can be seen.

viation. Implant listeners performed better on the task than did the four-band simulation listeners. However, no clear trend for increased masking with increasing duty cycle can be seen for either group. Instead, there is little change in performance for implant and simulation listeners from the condition of steady noise to any of the masker duty cycle conditions, especially for the +16-dB SNR condition. The fact that the SNR was quite favorable for the implant and simulation listeners (+8 and +16 dB), and the lack of change with duty cycle suggest that the effect of the noise may not be so much energetic masking of the signal, but informational masking, possibly caused by the fluctuations in the noise masker being misinterpreted as part of the signal. In contrast, the normal-hearing listeners showed clear energetic masking at −8- and −16-dB SNR.

The results of Experiment 1a demonstrate that the interference caused by gated noise maskers for implant (and simulation) listeners is not only related to energetic masking. For these stimuli and maskers, there is no systematic relationship between the percentage of time the masker is present and keyword recognition. Instead, the gated noise

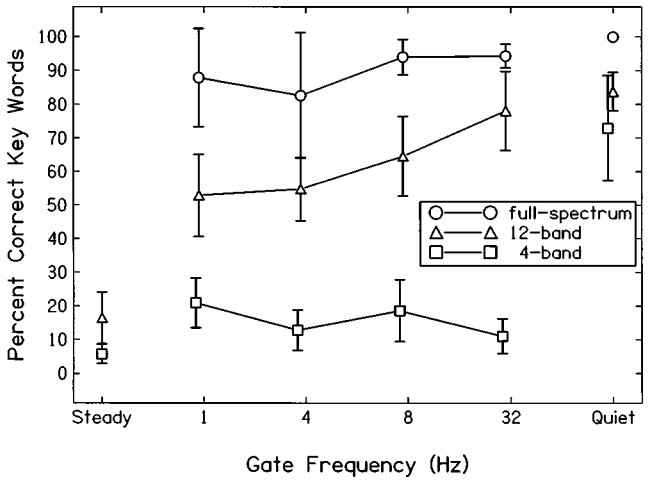


FIG. 5. Performance of normal-hearing listeners is shown for their understanding of sentences in gated noise when the stimuli were full-spectrum (open circles), 12-band simulations (open triangles), and 4-band simulations (open squares). Mean percent correct keywords (±1 s.d.) is plotted as a function of noise gate frequency. All stimuli were presented at −4-dB SNR.

masker used in Experiment 1 seems to disrupt the flow of information for simulation listeners and implant users.

B. Experiment 1b: Understanding speech in gated noise with increasing spectral channels

In Experiment 1b, normal-hearing listeners identified keywords from sentences in gated noise when the stimuli were full-spectrum, 12-band simulations, or 4-band simulations. Participants were eight additional normal-hearing college students who had not previously participated in any experiments. The duty cycle of the noise remained constant at 50%. It was hypothesized that the differences in spectral information present in the three sets of stimuli would have significant effects on listener performance. The results are shown in Fig. 5.

Figure 5 demonstrates the performance of normal-hearing listeners for their understanding of sentences in gated noise when the stimuli were full-spectrum (open circles), 12-band simulations (open triangles), and 4-band simulations (open squares). Mean percent correct keywords is shown as a function of noise gate frequency while the SNR remained constant at −4 dB. All three types of stimuli were highly intelligible in quiet (72% correct for 4 bands, 83% correct for 12 bands), but were almost completely unintelligible (5% and 12%, respectively) in steady noise.

The current listeners showed no release from masking in the four-band simulation condition, as was demonstrated previously in Nelson *et al.* (2003). In this condition, the identification of keywords in gated noise was not significantly different from identification in steady noise. The results from the 12-band simulations are quite different, however. Listeners showed significant release from masking when the stimuli had 12 bands of spectral information. A two-factor repeated measures ANOVA indicated a significant difference among spectral conditions (full-spectrum, 12-band and 4-band: $F[2,16]=409.5$, $p<0.0001$). There was also a significant difference among gate frequencies ($F[3,24]=14.6, p<0.0001$) and a significant gate frequency by spec-

trum interaction ($F[6,48]=9.3, p<0.01$). For the 12-band simulation stimuli, for which the results were not limited by ceiling or floor effects, the faster the gate frequency, the greater the release from masking. Performance for 12-band gated noise conditions improved (over steady noise) by approximately 40% for slow gate frequencies, and by as much as 60% for the fastest gate frequency (32 Hz). These results suggest that sentence stimuli with greater spectral detail are more easily segregated from gated background noise than are stimuli with poor spectral representation.

The results of Experiment 1 suggest that when implant listeners attempt to understand speech in gated noise, the fluctuations in the noise appear disruptive and confusing, more than energetically masking. Performance of implant and simulation listeners was not significantly better for the 50% duty cycle condition compared to the 75% condition, which would be expected if the disruption were due to the energy of the noise. In addition, it is presumed that energetic masking at +16-dB SNR should be minimal. We hypothesize that the envelope disruptions reduce masking release at least partially because there is not a clear segregation of the speech signal from the noise. As Qin and Oxenham (2003) hypothesize, the weak representation of pitch may contribute to the lack of segregation. As a result, implant and simulation listeners are not able to listen in the dips of the noise, not because the target speech in the dips is inaudible when the speech is 65 dBA and the noise 49 dBA, but because listeners do not group consecutive glimpses of the quiet signal. Masking release is facilitated by improving the spectral representation of the speech signal, presumably because the greater spectral representation of the signal allows more successful grouping of the speech signals and segregation of speech from noise.

C. Experiment 2a: Understanding gated sentences

In Experiment 2, listeners identified sentences that were themselves gated, or interrupted by periods of silence. Identifying gated speech was considered to be a direct measure of auditory fusion of the speech signal. The more that listeners are able to fuse the interrupted speech signal into a coherent stream, the more successful they may be at identifying key words in gated sentences.

Figure 6 shows the performance of normal-hearing listeners (open squares), implant listeners (open circles) and four-band simulation listeners (open triangles) for understanding gated sentences in quiet. Mean percent keywords correct is shown as a function of the gate frequency. As a reference point, recall from Figs. 2 and 3 that the quiet performance of the implant listeners was approximately 80% correct, and for the simulation listeners' quiet performance was approximately 60% correct.

Normal-hearing listeners showed improved performance with increasing gate frequencies above about 4 Hz, with significantly better performance at 16 and 32 Hz than at 2 and 4 Hz. Their identification of gated sentences at 16 and 32 Hz gate frequency was very good, approximately 80% correct. Thus, original full-spectrum speech stimuli that were interrupted with brief periods of silence were easily fused into an intelligible sentence.

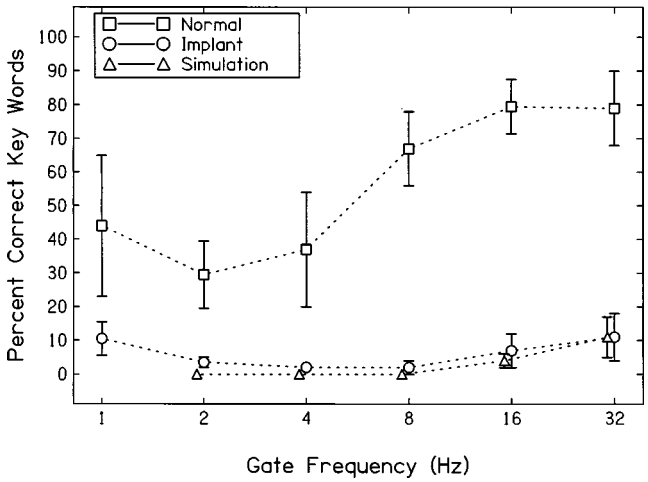


FIG. 6. Performance of normal-hearing listeners (open squares), implant listeners (open circles), and four-band simulation listeners (open triangles) for understanding gated sentences in quiet. Mean percent keywords correct (±1 s.d.) is shown as a function of the gate frequency.

In contrast, implant and four-band simulation listeners performed surprisingly poorly on this task at all gate frequencies. There was no significant difference between implant and the four-band simulation listeners on this task. They rarely scored above 10% correct for any condition, suggesting that they did not tolerate interruptions of any kind in the continuous speech signal. They showed little evidence of fusing across the brief silent periods. Presumably, the lack of fusing might be due to the sparse spectral information available to the implant and four-band simulation listeners. In addition, it may be because the periods of silence disrupted the temporal envelope of the speech signal.

D. Experiment 2b: Understanding gated sentences with increasing spectral channels

Experiment 2b tested the hypothesis that improving the spectral representation of speech would improve the simulation listeners' ability to fuse (and consequently understand) interrupted speech. Figure 7 shows the performance of eight additional normal-hearing listeners for their understanding of gated sentences when the stimuli were full-spectrum (open circles), 12-band simulations (open triangles), and four-band simulations (open squares). Mean percent correct keywords is shown as a function of sentence gate frequency.

A two-factor repeated measures ANOVA indicated a significant difference among spectral conditions (full-spectrum, 12-band and 4-band: $F[2,16]=452.4$, $p<0.0001$). There was also a significant difference among gate frequencies ($F[3,24]=181.5, p<0.0001$) and a significant gate frequency by spectrum interaction ($F[6,48]=76.6, p<0.0001$). At faster gate frequencies, interrupted speech recognition improved for the 12-band and full-spectrum stimuli.

As in Experiment 2a, the listeners in this experiment showed improved keyword recognition with increasing gate frequency. The performance of listeners identifying full-spectrum stimuli interrupted by 32 Hz silent intervals was not significantly different from their identification of continuous sentences. Natural speech that was interrupted by brief silent gaps (16 ms) was highly intelligible, although

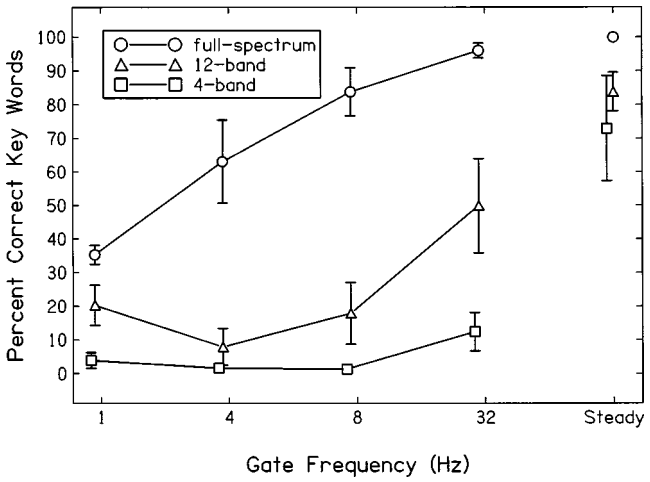


FIG. 7. Performance of normal-hearing listeners is shown for gated sentences when the stimuli were full-spectrum (open circles), 12-band simulations (open triangles), and 4-band simulations (open squares). Mean percent correct keywords (±1 s.d.) is shown as a function of sentence gate frequency.

listeners reported that it sounded ''rough.'' In contrast, speech that was interrupted by longer silent gaps had missing phonemes and syllables that caused a reduction in intelligibility. Performance for the four-band stimuli in this study was very poor and was equivalent to that reported in Experiment 2a above. Simulation listeners with only four bands of modulated noise were unable to understand more than about 10% of keywords in interrupted sentences at any gate frequency. Listener performance for the 12-band simulation condition was significantly better than for the 4-band conditions, and was significantly poorer than the full-spectrum condition. This result implies that spectral information is important for fusing interrupted sentences. As the amount of spectral information increases, listeners improve in their ability to integrate across silent intervals in gated speech.

IV. GENERAL DISCUSSION

Results from this investigation and from the previous related work (Nelson *et al.*, 2003) suggest that cochlear implant users are significantly and negatively affected by fluctuating signals and by fluctuating background noise. Even when temporal gaps in noise are as long as 500 ms, listeners do not show significant release from masking when compared to their performance in steady noise. Background noises as low as +16-dB SNR cause significant disruption in speech perception. The amount of masking caused by fluctuating noise seems to be only weakly related to the modulation frequency or the duty cycle of the noise. The fluctuating nature of the noise appears to serve as a disruption and an informational masker, rather than as an energetic masker.

Implant users not only show limited improvement from temporal gaps in fluctuating maskers, they also do not demonstrate successful auditory fusion of interrupted speech signals. Normal-hearing listeners presented with natural speech show improving word recognition performance with increasing rate of fluctuation. Implant and simulation listeners did not. It has been proposed by McKay and Carlyon (1999) and others (e.g., Guerts and Wouters, 2001; Xu, Tsai, and Pfingst, 2002; Green, Faulkner and Rosen, 2002) that implants have a relatively weak temporal pitch that does not convey voice fundamental frequency in a robust manner. The current results are consistent with the idea that implant listeners' perception of voice pitch is not strong enough to aid in segregating speech from noise, or in fusing interrupted speech. Implant and simulation listeners do not appear to successfully segregate speech signals from gated background noise, and do not fuse interrupted speech signals into a coherent speech stream.

The nature of the impairment seems to be that implant listeners receive a signal with reduced spectral cues, which then results in greater perceptual importance of the amplitude envelope of incoming signals for word and sentence identification (as in Hedrick and Carney, 1997). The envelopes of the signals, however, are affected by the fluctuations introduced by gating the speech or the background noise, and thus they are unreliable cues. Improving the spectral representation of speech, as shown in Experiments 1b and 2b and in the work of Qin and Oxenham (2003), seems to improve simulation listeners' understanding of sentences, and therefore improves their resilience to fluctuations in background noise or speech. Increasing the number of simulation channels does not, however, restore performance to that obtained with natural speech. If improved spectral resolution is important for implant listeners, then the ''best'' implant listeners would presumably be those who have better spectral (spatial) resolution, better word recognition in quiet, and better masking release and auditory fusion of interrupted sentences. This issue warrants further investigation. It seems logical that those listeners with the best spectral resolution would have the highest word identification scores (as in Henry and Turner, 2003), but it is unknown whether those better listeners have the greatest masking release and/or best identification of gated sentences. To test the hypothesis at a first approximation, correlations between the keyword identification score (percent correct in quiet) and amount of masking release (expressed as the difference between performance in gated and continuous noise in percent correct) were calculated. Similarly, we obtained correlations between percent correct words in quiet and auditory fusion (calculated by the average percent correct gated sentence score). We observed a moderate correlation ($r=0.5$) between keywords correct in quiet and masking release for these nine cochlear implant users, and no apparent correlation ($r=0.1$) between keywords correct in quiet and gated sentence performance. It may be, as Qin and Oxenham (2003) hypothesized, that improved spectral representation of speech signals is not sufficient for full masking release and/or auditory fusion of interrupted speech. Instead, they hypothesize that a stronger representation of voice pitch may be necessary for implant and simulation listeners to approach normal performance on such tasks. An additional investigation of the role of spectral resolution versus voice pitch strength is needed.

Zeng (2002) has proposed future cochlear implant-algorithms that might improve implant listeners' pitch resolution and range. The work of Qin and Oxenham (2003) supports the premise that an improved representation of F0 may help implant users significantly. Those proposed poten-

tial solutions include: increasing the physical number of electrodes, introducing virtual channels between electrodes (e.g., McDermott and McKay, 1994), and restoring stochastic electric responses (e.g., Zeng, Fu, and Morse, 2000). The current findings support the premise that improvements in implant processing could improve implant listeners' masking release and auditory fusion, provided that the underlying (neural) spatial resolution is adequate.

V. SUMMARY AND CONCLUSIONS

Cochlear implant users are significantly affected by fluctuating signals and by fluctuating background noise. Based on the results of Experiment 1, we conclude that implant and simulation listeners show limited improvement from temporal gaps in fluctuating maskers, and do not demonstrate successful auditory fusion of interrupted speech signals. Implant listeners apparently receive a signal with reduced spectral information that results in listeners' reliance on fluctuations in the amplitude of incoming signals. However, the amplitude of the signal is affected by the fluctuations introduced by gating the speech or the background noise. Based on Experiment 2, we conclude that implant and simulation listeners do not appear to successfully fuse interrupted speech signals into a coherent speech stream. Improving the spectral representation of speech seems to improve listeners' stream segregation and resilience to fluctuations in background noise or speech. Speech processing strategies that convey voice pitch more successfully may be better for providing masking release and fusion of interrupted speech.

ACKNOWLEDGMENTS

This research was supported by NIH Grant No. P50-DC00110. The collegial help of Arlene Carney and Gail Donaldson, and the work of research assistants Brian Urban and Laura Price are gratefully appreciated. The helpful comments of two anonymous reviewers provided significant improvements to the manuscript.

[1]The quiet speech signals were processed and later mixed with noise. We considered mixing the speech and noise prior to processing during the early stages of our experiments, but found that the number of stimuli (60 lists of 10 talkers) and conditions (five SNRs, three sets of simulation channels, three duty cycles) was unwieldy for preprocessing and storage of processed stimuli. In addition, the fact that the noise would be ''frozen'' for each stimulus was undesirable. As a comparison, we mixed a few sentences from several talkers with gated noise at +8-dB SNR and processed the mixed stimuli. Those stimuli were compared visually and via listening with those that were used (the processed sentences plus noise.) We could not detect a difference between the stimuli, and the results of pilot listening showed very similar recognition. In addition, the results obtained from the experiments are very similar to other reports from the literature regarding implant simulations. Results from the four-band simulations are quite similar to those from our actual implant listeners, as has also been noted by others. As such, we decided to process all the sentences prior to mixing them with noise in real time. We recognize that this is not exactly what implant listeners would experience, but believe that the results are revealing about the nature of implant listening in noise.

Bacon, S. P., Opie, J. M., and Montoya, D. Y. (**1998**). ''The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds,'' J. Speech Lang. Hear. Res. **41**, 549–563.

Bashford, J. A., and Warren, R. M. (**1987**). ''Multiple phonemic restoriations follow the rules for auditory induction,'' Percept. Psychophys. **42**, 114–121.

Chatterjee, M., and Galvin III, J. J. (**2002**). ''Auditory streaming in cochlear implant listeners,'' J. Acoust. Soc. Am. **111**, 2429.

Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., and Shunn-Cunningham, B. G. (**2003**). ''Note on informational masking,'' J. Acoust. Soc. Am. **113**, 2984–2987.

Green, T., Faulkner, A., and Rosen, S. (**2002**). ''Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants,'' J. Acoust. Soc. Am. **112**, 2155–2164.

Grimault, N., Bacon, S., and Micheyl, C. (**2002**). ''Auditory stream segregation on the basis of amplitude-modulation rate,'' J. Acoust. Soc. Am. **111**, 1340–1348.

Guerts, L., and Wouters, J. (**2001**). ''Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants,'' J. Acoust. Soc. Am. **109**, 713–726.

Gustafsson, H. A., and Arlinger, S. D. (**1994**). ''Masking of speech by amplitude-modulated noise,'' J. Acoust. Soc. Am. **95**, 518–529.

Hedrick, M. S., and Carney, A. E. (**1997**). ''Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants,'' J. Speech Lang. Hear. Res. **40**, 1445–1457.

Hedrick, M. S., and Jesteadt, W. (**1996**). ''Effect of relative amplitude, presentation level, and vowel duration on perception of voiceless stop consonants by normal and hearing-impaired listeners,'' J. Acoust. Soc. Am. **100**, 3398–3407.

Henry, B. A., and Turner, C. W. (**2003**). ''The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners,'' J. Acoust. Soc. Am. **113**, 2861–2873.

IEEE (**1969**). ''IEEE recommended practice for speech quality measurements,'' IEEE Trans. Audio Electroacoust. **17**, 225–246.

Kwon, B. J., and Turner, C. W. (**2001**). ''Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference,'' J. Acoust. Soc. Am. **110**, 1130–1140.

McDermott, H. J., and McKay, C. M. (**1994**). ''Pitch ranking with nonsimultaneous dual-electrode stimulation of the cochlea,'' J. Acoust. Soc. Am. **101**, 1622–1631.

McKay, C. M., and Carlyon, R. P. (**1999**). ''Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains,'' J. Acoust. Soc. Am. **105**, 347–357.

Miller, G. A., and Licklider, J. C. R. (**1950**). ''The intelligibility of interrupted speech,'' J. Acoust. Soc. Am. **19**, 167–173.

Nelson, P., Jin, S.-H., Carney, A. E., and Nelson, D. A. (**2003**). ''Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners,'' J. Acoust. Soc. Am. **113**, 961–968.

Punch, J. L. (**1978**). ''Masking of spondees by interrupted noise in hearing-impaired listeners,'' J. Am. Audiol. Soc. **3**, 245–252.

Qin, M. K., and Oxenham, A. J. (**2003**). ''Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers,'' J. Acoust. Soc. Am. **114**, 446–454.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). ''Speech perception without spectral cues,'' Science **270**, 303–304.

Warren, R. W. (**1999**). *Auditory Perception: A New Analysis and Synthesis* (Cambridge University Press, New York).

Wightman, F. L., Callahan, M. R., Lutfi, R. A., Kistler, D. J., and Oh, E. (**2003**). ''Children's detection of pure-tone signals: Informational masking with contralateral maskers,'' J. Acoust. Soc. Am. **113**, 3297–3305.

Xu, L., Tsai, Y., and Pfingst, B. (**2002**). ''Features of stimulation affecting tonal-speech perception: Implications for cochlear prostheses,'' J. Acoust. Soc. Am. **112**, 247–258.

Zeng, F. G. (**2002**). ''Temporal pitch in electric hearing,'' Hear. Res. **174**, 101–106.

Zeng, F. G., Fu, Q. J., and Morse, R. (**2000**). ''Human hearing enhanced by noise,'' Brain Res. **869**, 251–255.

Independent component analysis for automatic note extraction from musical trills[a)]

Judith C. Brown[b)]
Physics Department, Wellesley College, Wellesley, Massachusetts 02181 and Media Lab, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Paris Smaragdis[c)]
Mitsubishi Electric Research Lab, Cambridge, Massachusetts 02139

(Received 19 June 2003; accepted for publication 9 February 2004)

The method of principal component analysis, which is based on second-order statistics (or linear independence), has long been used for redundancy reduction of audio data. The more recent technique of independent component analysis, enforcing much stricter statistical criteria based on higher-order statistical independence, is introduced and shown to be far superior in separating independent musical sources. This theory has been applied to piano trills and a database of trill rates was assembled from experiments with a computer-driven piano, recordings of a professional pianist, and commercially available compact disks. The method of independent component analysis has thus been shown to be an outstanding, effective means of automatically extracting interesting musical information from a sea of redundant data. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1698774]

PACS numbers: 43.75.St, 43.60.Cg, 43.75.Mn [SEM] Pages: 2295–2306

I. INTRODUCTION

As with many fields today the processing of digitized musical information is inundated with huge masses of data, much of which is redundant. One attempt to reduce this deluge of data in the musical domain was attempted a decade ago with principal component analysis or PCA (Stapleton and Bass, 1988; Sandell and Martens, 1995). Earlier Kramer and Mathews (1956) had written an excellent introduction to data reduction in audio.

Since then the field of information processing has made a stride forward with new algorithms, one of particular interest to audio being independent component analysis [ICA; see Hyvarinen (1999) for an excellent introduction]. This method has been used with success in what is called blind source separation (Torkkola, 1999) and is a solution under certain restrictions to the computational statement of the age old cocktail party effect, addressing the question of whether a machine can emulate a human in picking out a single voice in the presence of other sources. The solution to this problem is considered by many to be the holy grail of audio signal processing. A restriction in the mainstream use of ICA has been that the number of microphones must be equal to or greater than the number of sources.

Recent reports (Casey and Westner, 2000; Smaragdis, 2001; Brown and Smaragdis, 2002) have indicated that, if a signal is preprocessed into frames of magnitude spectral features, then independent component analysis can be applied without the constraint of multiple microphones to extract the features carrying maximum information. We will develop this method further for the analysis of trills with a twofold purpose:

(1) Automatic redundancy reduction—We will show that ICA can be used to obtain musical information quickly, easily, and accurately from data recorded with a single microphone.
(2) Creation of database—From the calculations with ICA, we will assemble a database of information on a large number of trills obtained from a variety of sources to draw conclusions about trill rates.

II. BACKGROUND

A. Statistics background

Most of the sensory information we receive is highly redundant, and the goal of acoustical signal processing is often to expose the fundamental information and disregard redundant data. Since this is a common problem in data processing, statistical methods have been devised to deal with it. The following sections describe two of the most powerful techniques applied to spectral audio data.

1. Principal component analysis

A number of data reduction techniques are based on finding eigenfunctions for the second-order statistics of the data (Therrien, 1989). These techniques attempt to approximate a given data set using the superposition of a set of linearly independent functions, called basis functions, in a manner similar to the approximation of a sound by the superposition of sinusoids. Using a number of basis functions that equals the dimensionality of the original data set gives a perfect reconstruction. More often, the use of a reduced set of these functions results in efficient data encoding or a more useful interpretation of the data. The most prominent of these

[a)]Portions of these results were first presented at the 143rd ASA Meeting in Pittsburgh, PA (Brown and Smaragdis, 2002).
[b)]Electronic mail: brown@media.mit.edu
[c)]Electronic mail: paris@merl.com

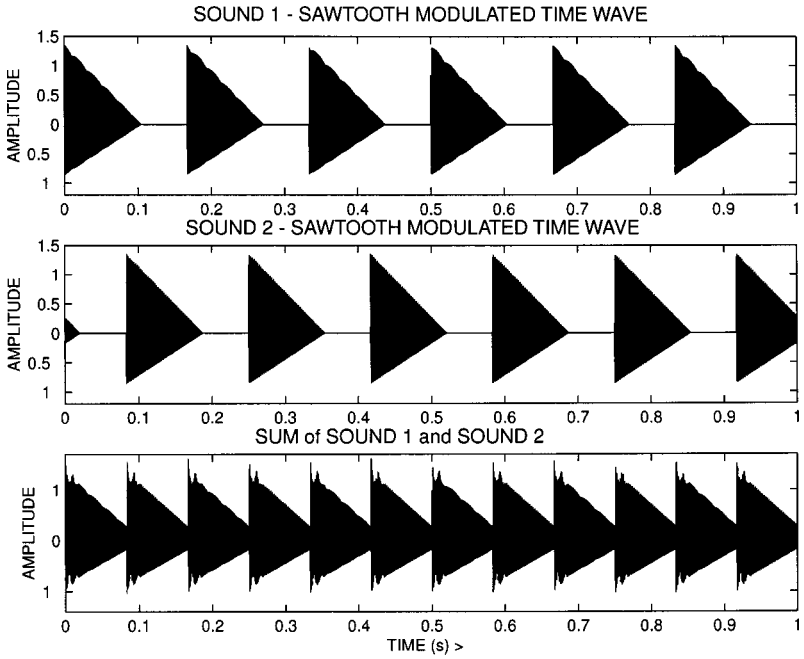


FIG. 1. Synthetic signal simulating a trill and consisting of the sum of two complex sounds, each containing three harmonics and modulated by a low-frequency sawtooth. The upper two graphs are the individual complex sounds, and the bottom graph is the sum.

approaches is called principal component analysis in the statistics literature, also referred to as the Karhunen–Loeve transform in the signal processing literature.

More formally, given a set of data vectors of dimension N, the method of principal component analysis can be used to find new a set of N' ($N' \leqslant N$) basis functions which are uncorrelated (second-order independence) and can be used to reconstruct the input. These are optimal in the sense that no other set of N' vectors gives a better least mean squares fit. The new basis functions can be sorted by magnitude of their variance, which is a measure of their importance in describing the data set. Optionally we can ignore the least important bases, and the dimensionality of the data set can be reduced with fine detail eliminated.

As an example applicable to our later sections, we consider the matrix of values calculated for the magnitude of the constant-Q transform (Fourier transform with log-frequency spacing) of a temporal waveform broken up into N shorter time segments. The calculation was carried out by the method of Brown (1991; Brown and Puckette, 1992) with a Q of 17 corresponding to the frequencies of musical notes. The time wave is the sum of two synthetic sounds with fundamental frequencies corresponding to musical notes C_6 and D_6 and each containing harmonics two and three. These sounds are amplitude modulated by a low-frequency sawtooth simulating alternating notes as found in trills (Fig. 1).

Figure 2 is a plot of the constant-Q coefficients calculated for the input time wave of Fig. 1. Each column repre-

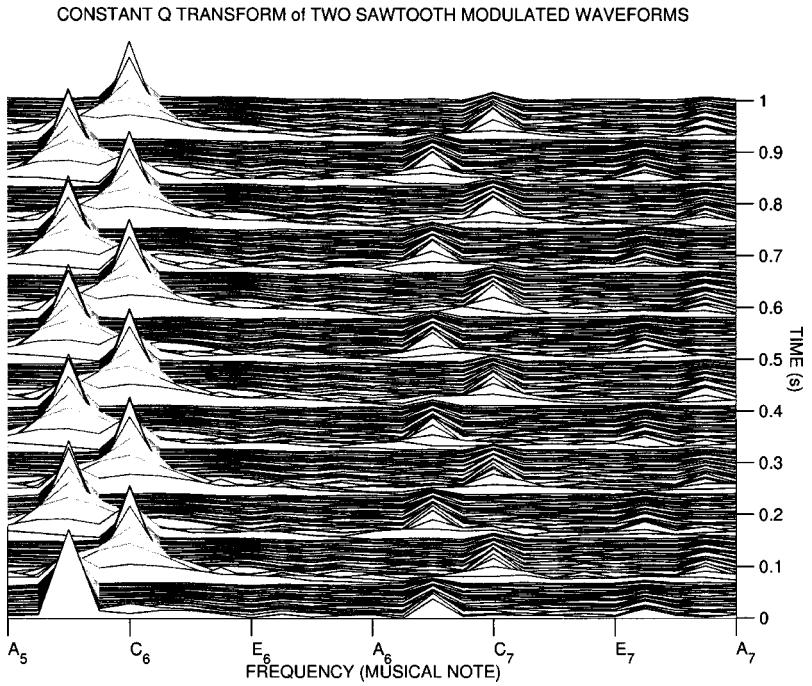


FIG. 2. Magnitude (arbitrary units) of the constant-Q transform against frequency and time in seconds (waterfall plot) for the complex sound of Fig. 1. Frequencies are indicated on the horizontal axis by musical notes.

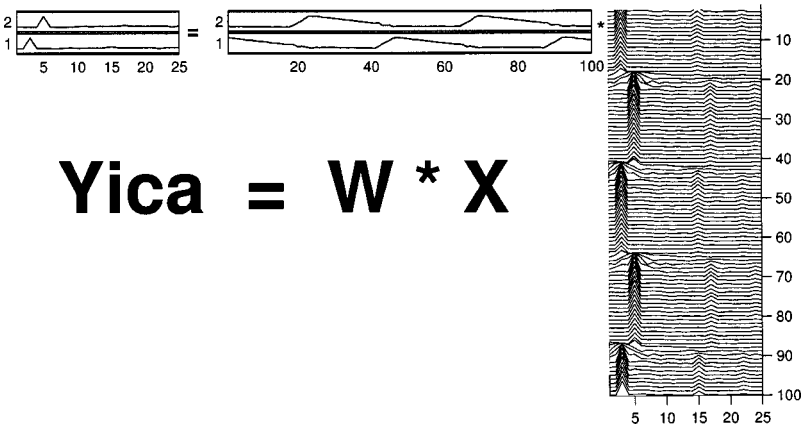


FIG. 3. Graphical example of the matrix mulitplication $Y=W*X$ for the first third of the data matrix of Fig. 2, keeping the two most important independent components. The independent components **Y** for this orientation of the data matrix are the frequency bases. Note that the two basic shapes of the rows of X have been extracted. The transformation matrix **W** displays the temporal behavior of the two independent components (same shape as columns three and five) and are referred to as time bases.

sents the values of one spectral coefficient at N times, and each row consists of M frequency samples of a single variate. Viewed as a whole, the columns are components of a random vector, and each column is a sample of that vector at a different frequency. These data are highly redundant with one basic shape for the spectra of the two notes present differing only in their horizontal positions.

It is more common to consider the transpose of this matrix, which gives samples in time for the rows, but better results were obtained as described. This is because the frequency-dependent rows (or samples) are better separated and hence less correlated for the covariance calculation.

Subtracting the average of each row from the elements of that row, and defining a typical element of the covariance matrix **C** (Therrien, 1989) as the expectation value, we have

$$C_{ij}=\langle X_i \cdot X_j \rangle, \tag{1}$$

where the average is taken over all samples. See Appendix A for an example of this implementation.

For a finite data set where all samples are available as rows in a matrix X, the covariance matrix can be computed by

$$\mathbf{C}=\langle \mathbf{X}\cdot\mathbf{X^T} \rangle, \tag{2}$$

where $\mathbf{X^T}$ is the transpose of **X**. This matrix can be diagonalized by finding the unitary transformation **U** such that

$$\mathbf{U^T}\cdot\mathbf{C}\cdot\mathbf{U}=\mathbf{D}, \tag{3}$$

where **D** is diagonal. This is done by solving for the eigenvalues of **C** with the result

$$\mathbf{U^T}\cdot(\mathbf{X}\cdot\mathbf{X^T})\cdot\mathbf{U}=\mathbf{D}. \tag{4}$$

From Eq. (4) using the associative property of matrices and the transpose of a product

$$(\mathbf{U^T}\cdot\mathbf{X})\cdot(\mathbf{X^T}\cdot\mathbf{U})=(\mathbf{U^T}\cdot\mathbf{X})\cdot(\mathbf{U^T}\cdot\mathbf{X})^{\mathbf{T}}. \tag{5}$$

Defining a new matrix in Eq. (5),

$$\mathbf{Y_{pca}}=\mathbf{U^T}\cdot\mathbf{X}. \tag{6}$$

$\mathbf{Y_{pca}}$ is the matrix of principal components (called scores in the statistics literature) and has a diagonal covariance matrix with elements equal to the variances of its components. $\mathbf{U^T}$ is called the weights matrix in the statistics literature. Both $\mathbf{Y_{pca}}$ and the transformation matrix $\mathbf{U^T}$ can be ordered by magnitude of the variance. The dimensionality can thus be reduced by taking the k rows of each of these matrices corresponding to the largest variances. See Fig. 3 for an example of matrix multiplication keeping two components.

With this orientation of the data matrix X, the rows of $\mathbf{Y_{pca}}$ will be spectra corresponding to the rows with the largest variances and will be referred to as frequency bases. See the frequency dependence for the two complex sounds in Fig. 2. The rows of $\mathbf{U^T}$, the unitary transformation matrix, will show the time dependence for the k most important rows and be referred to as time bases.

Since the covariance matrix of $\mathbf{Y_{pca}}$ is diagonal, off-diagonal elements

$$D_{ij}=\langle Y_i \cdot Y_j \rangle \tag{7}$$

are zero showing that the components of **Y** are orthogonal or linearly independent. From a statistical point of view they are decorrelated showing

$$E[Y_iY_j]=E[Y_i]\cdot E[Y_j]=0. \tag{8}$$

This form of independence does not, however, mean that the two components are completely uncoupled and that they are statistically independent. For true statistical independence the joint probability density must factor into the marginal densities

$$p(Y_i,Y_j)=p(Y_i)\cdot p(Y_j), \tag{9}$$

and for this factorization to hold another method is needed.

2. Independent component analysis

The goal of independent component analysis is to find a linear transform

$$\mathbf{Y}=\mathbf{W}\cdot\mathbf{X} \tag{10}$$

such that the variates of Y are maximally independent. Stated otherwise, this transform should make the equation

$$p(Y_1,\ldots,Y_M)=\prod_{i=1}^{M} p(Y_i) \tag{11}$$

''as true as possible''. It is much more difficult to find the desired transformation **W** than the corresponding unitary transformation for PCA. One approach has been to minimize the relative entropy or Kuhlback–Liebler (KL) divergence (Deco and Obradovic, 1996). This is a quantity defined in information theory to give a measure of the difference in two

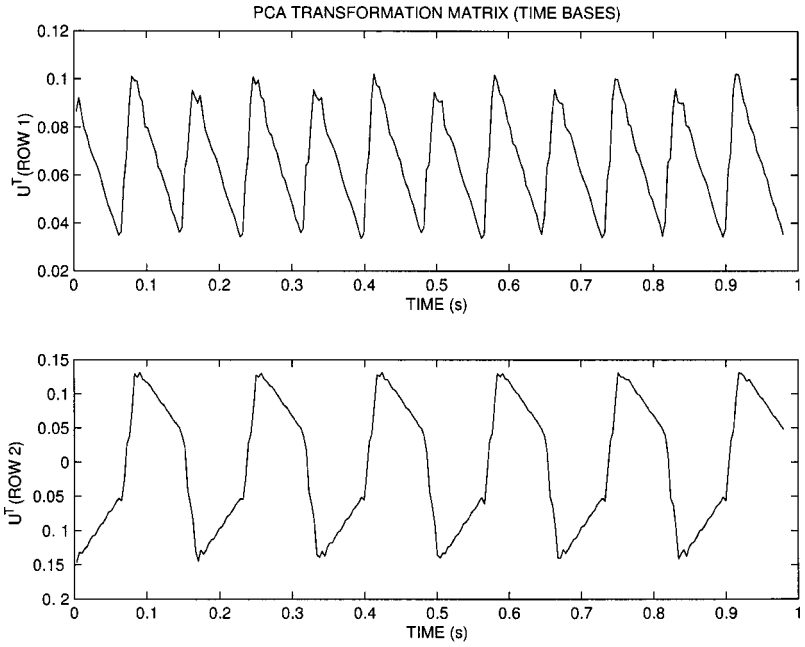


FIG. 4. PCA transformation matrix—the two most important rows of the unitary transformation matrix U^T of Eq. (6) for the complex sound of Fig. 1 with the constant-Q transform of Fig. 2 as data matrix X.

probability densities and has been used extensively for pattern classification.

The KL divergence is defined for two probability densities $p(x)$ and $q(x)$

$$K(p\|q)=\int p(x)\log\left(\frac{p(x)}{q(x)}\right)dx, \tag{12}$$

where the integral is taken over all x. The KL divergence can be easily adapted as a measure of the difference in the joint probability and the marginal densities in Eq. (11). In this context it is called the mutual information (Deco and Obradovic, 1996) $I(Y_1;\ldots;Y_M)$ and is a measure of the statistical independence of the variates whose densities appear on the right side of Eq. (11). That is, it tells us to what degree the Y_i are statistically independent:

$$I(Y_1;\ldots;Y_M)=K\left(p(Y_1,\ldots,Y_M)\left\|\prod_{i=1}^{M}p(Y_i)\right.\right). \tag{13}$$

Several algorithms for ICA solutions have used procedures which have the effect of minimizing the mutual information including those of Amari (1996) and Bell and Sejnowski (1995). These are called infomax and in general seek a transformation matrix **W** in Eq. (10) in an interative calculation.

An alternative approach, which is conceptually close to PCA, is to extend the second-order independence of PCA to higher orders using a cumulant-based method. This is the approach taken by Cardoso (1990; Cardoso and Souloumiac, 1996) in diagonalizing the quadricovariance tensor. Instead of the terms C_{ij} of the covariance matrix, he considers all products up to fourth order such as

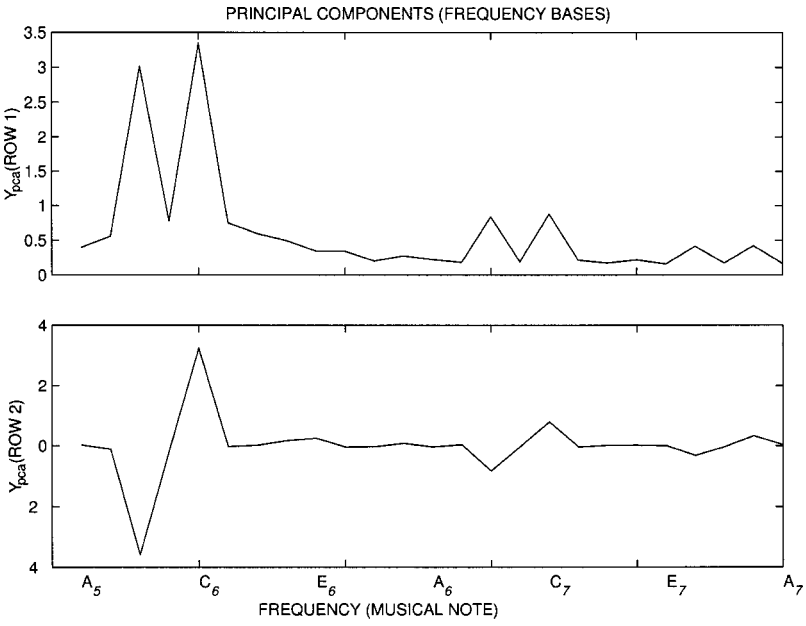


FIG. 5. The two most important principal components $\mathbf{Y_{pca}}$ from the transformation equation (6) for the complex sound of Fig. 1 with the constant-Q transform of Fig. 2 as data matrix X. Frequencies are indicated by musical notes on the horizontal axis. Note that the two basic shapes of Fig. 2 are mixed by the transformation.

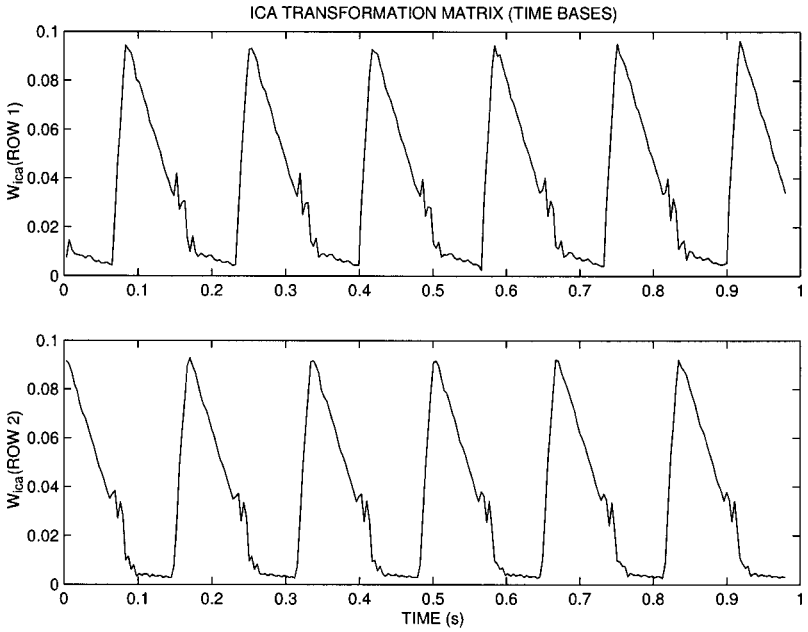


FIG. 6. ICA Transformation matrix—the two most important rows from the tranformation matrix **W** of Eq. (10) for the complex sound graphed in Fig. 1 with the constant-Q transform of Fig. 2 as data matrix *X*. This is also the first third of the matrix *W* in Fig. 3.

$$C_{ijkl} = \langle X_i \cdot X_j \cdot X_k \cdot X_l \rangle. \quad (14)$$

The diagonalization of this tensor ensures that no two dimensions of the data will have a statistical dependence up to and including the fourth order. This is a generalization of the diagonalization of the covariance matrix as done with PCA, where dependencies are eliminated up to second order. By extending the notion of the covariance matrix and forming the quadricovariance tensor (a fourth-order version of covariance), we effectively set a more stringent definition of statistical independence.

This concept can also be extended to an arbitrary order of independence by forming and diagonalizing even more complex structures. In this case the complexity of the process unfolds exponentially and can present computational issues. Fourth-order independence is a good compromise, exhibiting a manageable computational burden with good results.

B. Trill background

Trills were chosen for this study because they are extremely difficult to analyze. The note rate is very rapid, and when pedaled there are two temporally overlapping notes present. There is an advantage, however, in that they do not have simultaneous onsets.

The execution of trills has been studied by a number of groups interested either in performance on musical instruments or in perception limits of detection of two pure tones. The latter measurements are best summarized by Shonle and Horan (1976) who varied the frequency difference of two sinusoids with a modulation rate (frequency of a trill pair) of

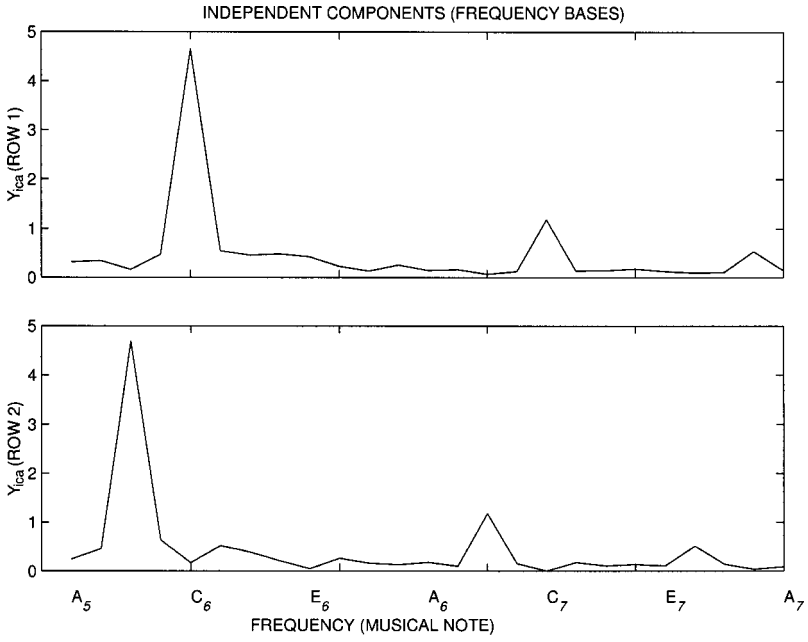


FIG. 7. Independent components—the two most important rows of the matrix **Y** of Eq. (10) for the complex sound graphed in Fig. 1 with transformation matrix **W** from the previous figure. See also Fig. 3. Frequencies are indicated by musical notes on the horizontal axis. It is clear that the calculation has picked up the 2nd harmonic 12 bins (an octave) above the fundamental and the 3rd harmonic 7 bins (a musical fifth) above that for each of these independent components.

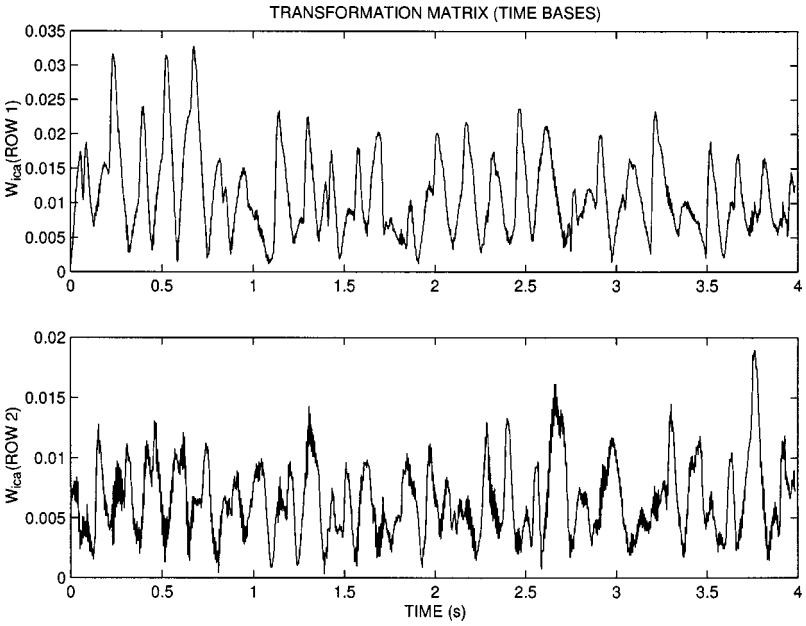


FIG. 8. ICA transformation matrix—the two most important rows from the matrix **W** of Eq. (10) for the constant-Q transform of the computer-driven Yamaha disklavier.

5 Hz and found that over the range 250–1000 Hz, fusion occurs at a difference frequency of roughly 30 Hz. Note that the modulation rate corresponds to a note rate of 10 Hz. The terminology ''note rate'' is used to avoid confusion with frequency of trill pairs. They conclude that a whole-tone trill (12% frequency difference) will be heard as alternating between two notes for frequencies over 400 Hz and as a warble below 125 Hz. The region between these frequencies is ambiguous and depends on the perception of the individual subject. See Table I for a comparison of these to other background studies.

Performance studies are more directly related to our results. Palmer (1996) found that the number of trills in an ornament depends on the tempo, which implies that the trill rate changes less than might otherwise be expected. Note rates varied from 11 Hz (measurement over 11 trill pairs) in a slow passage to 13.4 Hz (measurement over 9 trill pairs) in a fast passage.

Moore (1992) states that piano trills require one of the fastest alternating movements of which the hand is capable. He finds the upper limit to be about 12–14 notes/s.

In earlier work, Moore (1988) studied trills performed on a cello. He concluded that the limit on the trill seems to be derived from both the performer and the instrument. He gives no quantitative data, but his graphical data indicate a note rate of approximately 12 Hz.

III. SOUND DATABASE

The sounds analyzed consisted of two-note trills obtained from three sources:

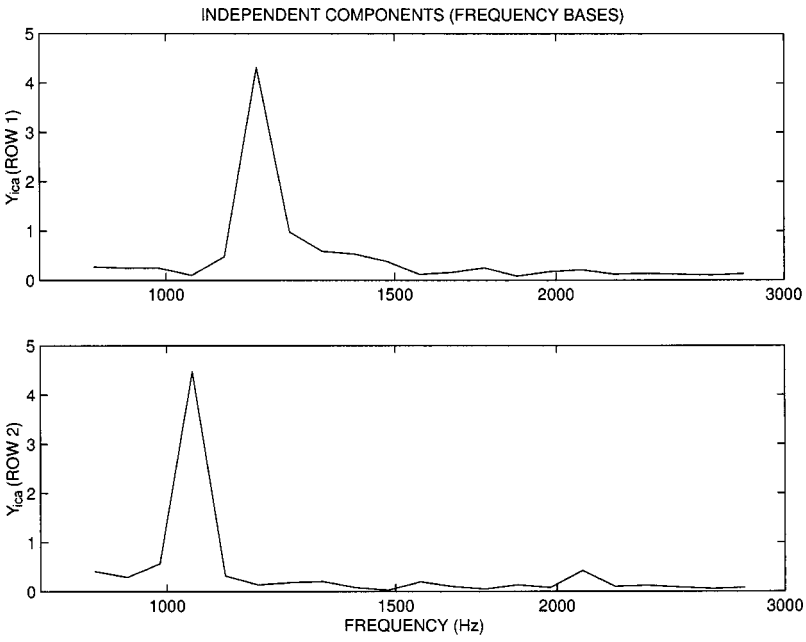


FIG. 9. Independent components—the two most important rows of the matrix **Y** of Eq. (10) for the computer-driven Yamaha disklavier with transformation matrix **W** from the previous figure.

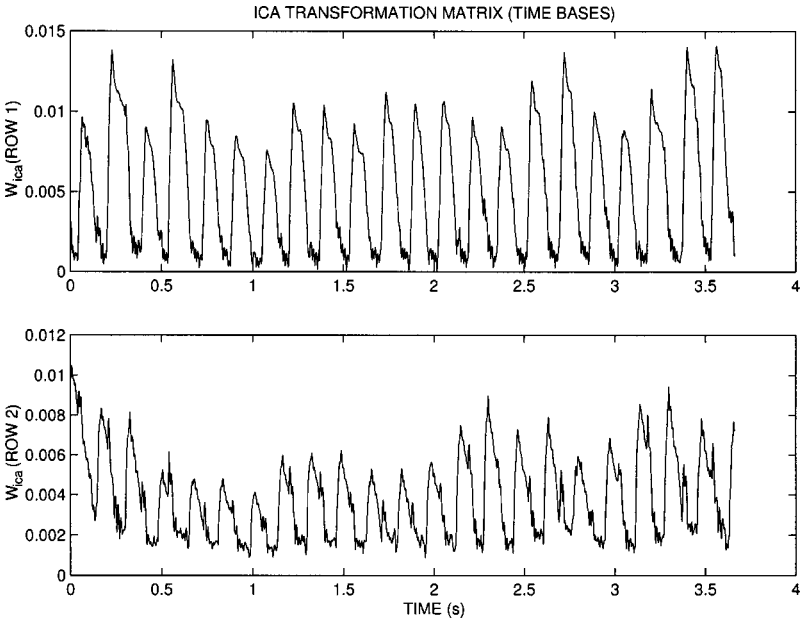


FIG. 10. ICA transformation matrix—the two most important rows from the matrix **W** of Eq. (10) for the constant-Q transform of a recorded performance by Charles Fisk. This is an example characterized as ''fast with control'' by the pianist.

(1) recordings of a Yamaha Disklavier piano programmed using Miller Puckette's ''pd'' program (Puckette, 1996) to drive the piano,
(2) recordings of pianist Charles Fisk of Wellesley College playing trills on a Steinway S, and
(3) excerpts from compact disks of performances by Ashkenazy, Horowitz, Goode, Wilde, and Pollini on piano, and Peter-Lukas Graf on the flute.

IV. CALCULATIONS AND RESULTS

Principal component analysis calculations were carried out using Matlab with the function *eig* for diagonalization of a matrix. See Appendix A for details. In our independent component analysis calculations (Appendix B), we used the algorithm Jade[1] and assumed that two notes were present by specifying two independent components in the calculation. If we assume fewer ICs than there are notes actually present, the independent components will consist of mixtures of the notes. If we assume more ICs than notes actually present, the notes will be evenly distributed across components.

A. Synthetic signal

Using known input as a first example, we compare the results using principal component analysis with those of independent component analysis for the computer-generated signal described in Figs. 1 and 2. Figures 4 and 5 show the quantities $\mathbf{U}^{\mathbf{T}}$, the transformation matrix, and $\mathbf{Y}_{\mathbf{pca}}$, the principal components, calculated from Eq. (6) and keeping the two most important principal components. The titles of the figures indicate frequency dependence (frequency basis functions) or time dependence (time basis functions).

Looking at the frequency bases of Fig. 5, we find that PCA has picked out the peaks corresponding to the two fundamental frequencies present. These are the dominant fre-

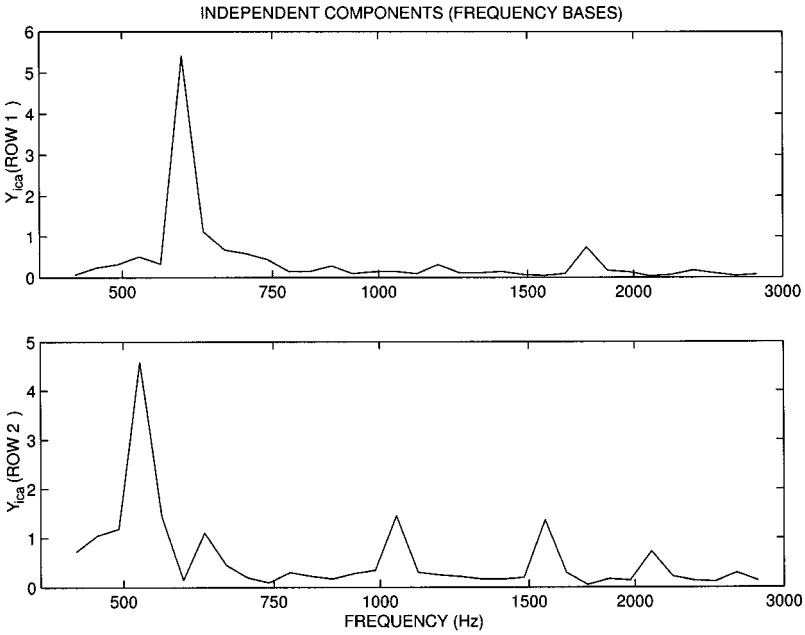


FIG. 11. Independent components—the two most important rows of the matrix **Y** of Eq. (10) for the recorded performance of Charles Fisk with transformation matrix **W** from the previous figure.

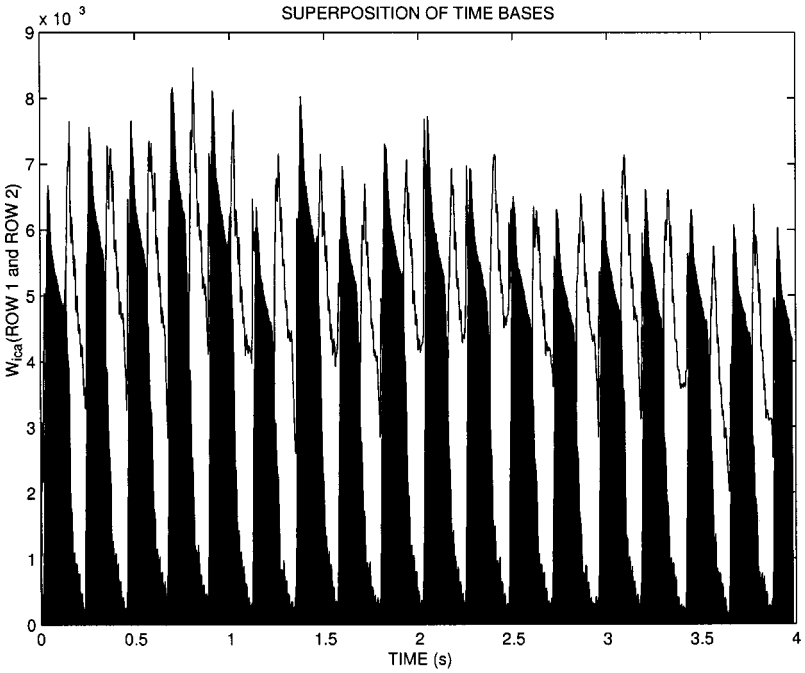


FIG. 12. Superposition of the times bases for one of the "slow" trills recorded by Charles Fisk. This shows clearly the spacing of the notes.

quencies in these data. But in choosing bases, PCA has chosen linear combinations of these two frequencies corresponding to the sum and difference of the two sources rather than separating them. This is a perfectly valid solution for PCA since these are orthogonal bases and are solutions to the eigenvalue equation Eq. (15).

Examining the time bases of Fig. 4 corresponding to these two principal components, they do not contain useful information about the temporal behavior of the two musical notes. The addition and subtraction has effectively removed the possibility of getting times of single note onsets.

Applying the ICA algorithm Jade to the same input (Fig. 2) to obtain **W** and **Y** of Eq. (10), the time bases and frequency bases seen in Figs. 6 and 7 are obtained. See also Fig. 3 for the operation applied to the first third of the file. Absolute values were plotted in these and other ICA results. The low-frequency sawtooth modulation of Fig. 6 is an excellent representation of the two alternating sounds simulating a trill, and the two independent components of Fig. 7 are a near-perfect extraction of the frequencies present in each of two complex sounds which were mixed. ICA has thus performed an excellent separation and yielded the two sources which are present while discarding redundant information.

B. Computer-driven piano

To test this method on real sounds, a Yamaha Disklavier piano was driven by computer at a number of different rates with whole-tone trills beginning on the notes C_5 or C_6. Recordings were made with a Sony TCD-D8 DAT recorder and

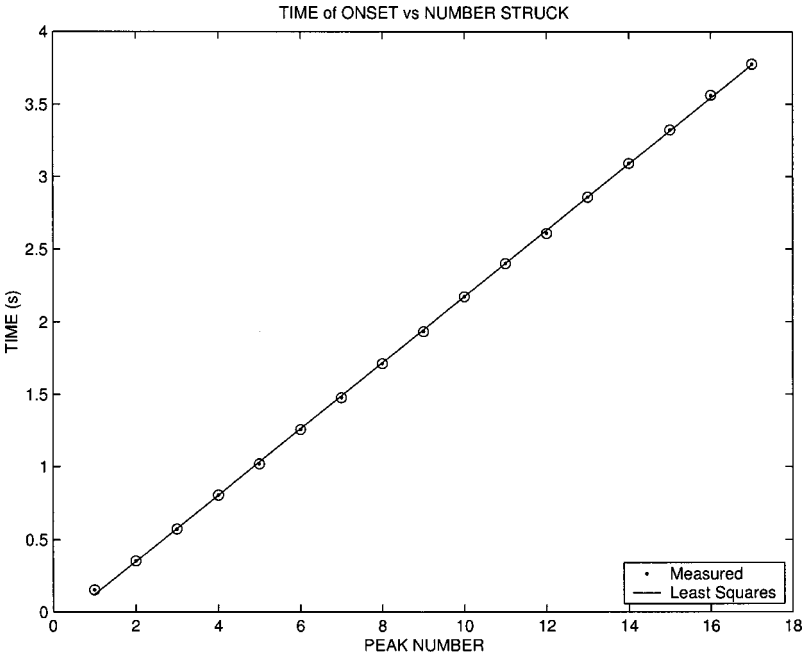


FIG. 13. Onset time against peak number for the peaks of one note of the previous figure compared to a least squares linear fit showing accuracy of note striking.

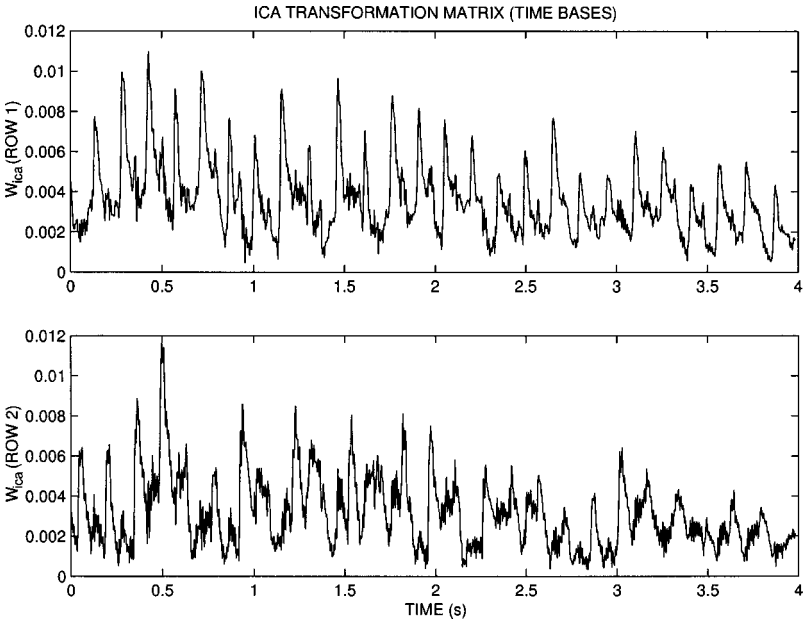


FIG. 14. Transformation matrix—the two most important rows from the matrix **W** of Eq. (10) for the constant-Q transform of the Pollini performance of Beethoven's Piano Sonata No. 32, Op. 111. This is included as an example of a performance analyzed from CD.

analyzed using the ICA algorithm Jade described previously. The example shown in Fig. 8 has a note rate of 13.5 Hz and is the maximum rate at which this piano could be driven without dropping notes; even so this example is not perfect for the time bases as it is a little beyond the region of reliable operation of the piano. The frequency bases of Fig. 9 are clearly separated, again demonstrating that ICA is able to pull out the relevant information while dropping redundant data.

C. Recordings of live performance

As an example of a live performance, Charles Fiske, a professional pianist and member of the performing faculty of the Wellesley College Music Department, generously agreed to do some trills for this study. In order to determine how a performer views trill rates, he was given the instructions to perform the trills slowly, fast with control, and very fast. These rates varied from 8.6 for slow to 12.1 notes/s for fast with control (Table I.) ICA results for the time bases and frequency bases are given in Figs. 10 and 11 for one of the ''fast with control'' examples.

Further analysis was carried out on one of the ''slow'' files and is shown in Figs. 12 and 13. The superposition of the time bases (black for one note, white for the other) is shown in Fig. 12 in order to demonstrate the precision of the alternating onsets. In a more quantitative graph, Fig. 13 shows the onset times for one of the two notes plotted against note number in order to obtain the average time between trill pairs. This is 0.22 s with a standard deviation of 0.01 showing that the trill is very precise.

D. Examples from compact disk

Trills from a number of performances on compact disk were studied since these had not been previously reported. In

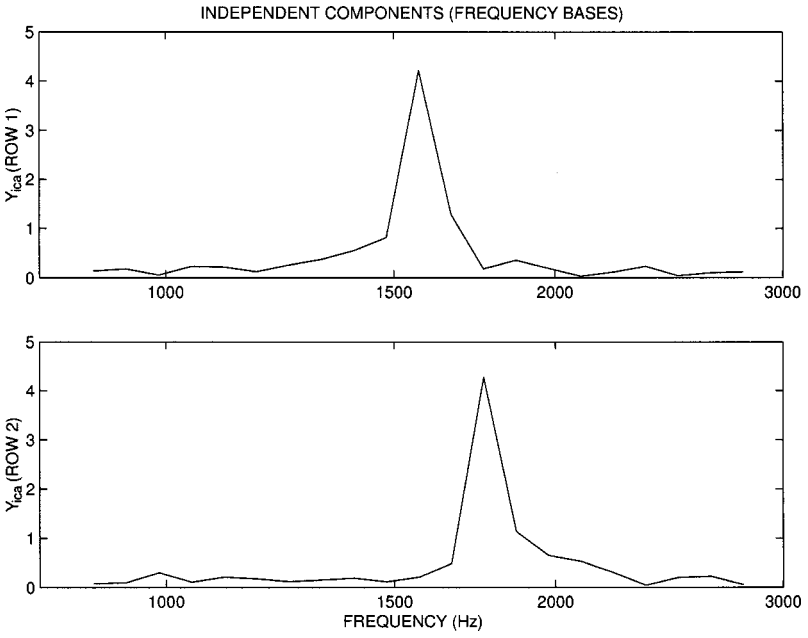


FIG. 15. Independent components—the two most important rows of the matrix **Y** of Eq. (10) for the Pollini performance from CD with transformation matrix **W** from the previous figure.

TABLE I. Summary of results on trill rates.

Reference or performer	Notes (or frequency)	Trill rate(note/s)	Comments
	Results from Literature		
Shonle/Horan		10, 16	from below and from above
Palmer		13.4, 11	fast passage, slow passage
Moore	D4 E4	13	
		12 to 14	upper limit
Michael Hawley		13	upper limit
	Computer-driven piano		
Yamaha 140	C6-D6	14.3	
Yamaha 150	C6-D6	13.5	
Yamaha 170	C6-D6	11.7	
	Recording of live performance		
Fisk	C5-D5	12.1	fast with control
Fisk	C5-D5	8.9	slow
Fisk	C6-D6	8.8, 8.6	slow—2 examples
	Performances from compact disk		
Pollini		13.5	
Ashkenazy CE 10		16.3	ornament
Ashkenazy MW	A4 B4	11.5	
Goode BA1	D5 E5	13.3	
Goode BA3	F5 G5	15	
Goode BW	G5 A5	13.2	
Horowitz BA1	D5 E5	11.2	
Horowitz BA2	E5f F5	13.8	
Horowitz BA3	F5 G5	13.3	
Horowitz BW	G5 A5	12.6	
Horowitz CE 10	C5 D5	15.7	ornament
Wild CE 10	C5 D5	16	ornament
Flute		12.8	

some cases difficulties in resolving the two notes were encountered due to pedaling, reverberation, or a significant difference in amplitudes of the two notes. Graphs of the transformation matrix and independent components for a particularly good example by Pollini playing Beethoven's Piano Sonata No. 32, Op. 111 are shown in Figs. 14 and 15. This is interesting in that the amplitudes of the two notes are almost exactly equal (arbitrary units on the vertical axis of Fig. 14), showing great control by the performer.

In order to demonstrate the applicability of this method to instruments other than the piano, our calculation was applied to a flute trill from Mozart's Flute Concerto No. 1. K313. The notes are extremely well resolved as seen in Fig. 16, but the amplitudes are not equal as in the previous ex-

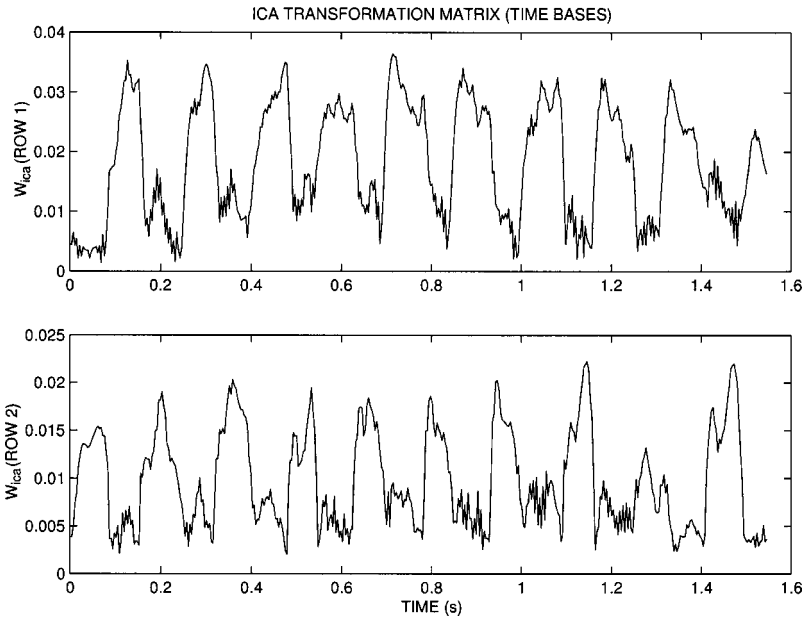


FIG. 16. ICA transformation matrix—the two most important rows of transformation matrix **W** of Eq. (10) for the constant-Q transform for the Mozart flute recording.

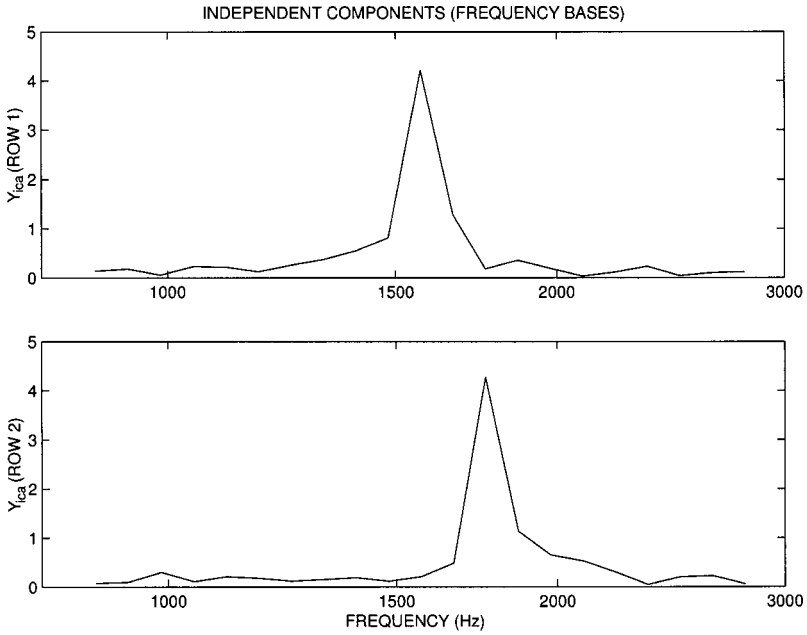


FIG. 17. Independent components—the two most important rows of the matrix **Y** of Eq. (10) for the Mozart flute recording with transformation matrix **W** from the previous figure.

ample by Pollini. The frequency bases from Fig. 17 show little evidence of higher harmonics indicating that in this frequency range the flute sound is close to a pure tone.

E. Summary of results on trills

Our data on trills is collected in Table I. Most of our results, including the flute trill, are in the range 12–14 notes/s predicted by Moore (1992) and pianist/computer scientist Michael Hawley[2] in a discussion with one of the authors. Pianist Charles Fisk in the recorded live performance was given the instructions to play slowly, and then fast with control. The ''fast with control'' example at 12.1 notes/s is consistent with Moore's and Hawley's predictions. The ornaments from the Chopin Etude Op. 10, No. 8 played by Ashkenazy, Horowitz, and Wild were all very fast at 16 notes/s, but this was not a sustained trill.

It is interesting to compare performances of the same trill by different performers. The first trill in Beethoven's Sonata Op. 57 (Appassionata) was played at 13.3 notes/s by Goode compared to 11.2 for Horowitz, which is significantly faster. The third trill in this piece was also significantly faster in the performance by Goode. And finally, a trill from the Beethoven's Sonata Op. 53 (Waldstein) offers a similar example. Thus there appears to be a consistent difference in the interpretations of these two performers. This opens a fertile area for further research in musical performance.

V. CONCLUSIONS

In this paper we have introduced a new method of musical analysis and applied it to musical trills. Redundancies inherent in the magnitude spectra of trills were identified, and statistical methods were employed to take advantage of this characteristic so as to reveal their basic structure.

The method of independent component analysis can simplify the description of trills as a set of note pairs described by their spectra and corresponding time envelopes. By examination of these pairs we can easily deduce the pitch and the timing of each note present in the trill. We have also noted how ICA, by employing higher-order statistics and forcing independence, improves the estimate compared to a straightforward application of principal component analysis.

The analysis itself is bootstrapped only to the data presented and devoid of any musical knowledge. In fact, it is a derivative of methods used for auditory scene analysis, which do not assume any previous auditory knowledge.[3] This fact allows us to analyze a wide variety of trills and not be constrained or biased by instrument selection, performance, or scale tuning issues. By avoiding the necessity of preprocessing for the extraction of semantically meaningful features, for example pitch or loudness, another advantage is found in a lower burden of computation and complexity.

Finally, we would like to stress the value of redundancy reduction for more complex musical analysis. We have shown how powerful this concept can be for trills; however, it is also applicable to more complex musical segments. In our future work we plan to expand upon this theme and demonstrate how this method can be applied to musical transcription.

APPENDIX A: IMPLEMENTATION OF PRINCIPAL COMPONENT ANALYSIS IN MATLAB

We take as input the matrix X, an N by M matrix in the orientation of Fig. 2. Using the function *eig* to find the unitary transformation U to diagonalize X,

$$[U,D]=eig(X*X'/M), \tag{A1}$$

where M is the number of columns of X. The eigenvalue matrix D is ordered by magnitude of its elements from low to high, and U is ordered correspondingly. If two principal components are desired, the last two columns of U will be taken and called the reduced matrix Ur. In matlab notation, $Ur=U(:,N-1:N)$. The transpose of Ur is plotted in the figures, and referred to as row 1 and row 2 in order of importance.

The function Y_{pca} is called the principal components in the figures and defined in Eq. (5) as

$$Y_{pca} = Ur' * X. \tag{A2}$$

APPENDIX B: IMPLEMENTATION OF INDEPENDENT COMPONENT ANALYSIS IN MATLAB

The algorithm *jade* was used for the independent component analyses in the form

$$[A, Yica] = jade(X, nc), \tag{B1}$$

where X is the matrix defined in Appendix A and nc is the number of components desired. $nc = 2$ in the calculations reported.

A is the inverse of the ICA transformation matrix so

$$W = pinv(A) \tag{B2}$$

and

$$Yica = W * X \tag{B3}$$

were plotted.

[1]A number of algorithms for performing independent component analysis are freely available on the internet, such as Jade, Amari, FastICA, and Bell. They can be found using an internet search engine, or more easily from links from Paris Smaragdis' home page. One file was checked with several of these algorithms to ensure that the results were independent of the algorithm used.

[2]Michael Hawley, personal communication.

[3]Because of this lack of knowledge many ICA-based algorithms are called ''blind.'' Knowledge accumulated from previous passes is not used and every example is treated as the first and only set of data the algorithm has encountered.

Amari, S., Cichocki, A., and Yang, H. H. (**1996**). ''A New Learning Algorithm for Blind Signal Separation,'' in Advances in Neural Information Processing Systems, edited by D. Touretzky, M. Mozer, and M. Hasselmo (MIT Press, Cambridge, MA).

Bell, A. J., and Sejnowski, T. J. (**1995**). ''An information-maximization approach to blind separation and blind deconvolution,'' Neural Comput. **7**, 1129–1159.

Brown, J. C. (**1991**). ''Calculation of a constant-Q spectral transform'' J. Acoust. Soc. Am. **89**, 425–434.

Brown, J. C., and Puckette, M. S. (**1992**). ''An efficient algorithm for the calculation of a constant-Q transform,'' J. Acoust. Soc. Am. **92**, 2698–2701.

Brown, J. C., and Smaragdis, P. (**2002**). ''Independent component analysis for onset detection in piano trills,'' J. Acoust. Soc. Am. **111**, 2446.

Cardoso, J. F. (**1990**). ''Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem,'' Proc. ICASSP, pp. 2655–2658.

Cardoso, J. F., and Souloumiac, A. (**1996**). ''Jacobi angles for simultaneous diagonalization,'' J. Math. Anal. Appl. **17**(1), 161–164.

Casey, M. A., and Westner, A. (**2000**). ''Separation of Mixed Audio Sources by Independent Subspace Analysis,'' in Proceedings of the International Computer Music Conference (ICMC), pp. 154–161.

Deco, G., and Obradovic, D. (**1996**). *An Information-theoretic Approach to Neural Computing* (Springer, New York, 1996).

Hyvarinen, A. (**1999**). ''Survey on Independent Component Analysis,'' Neural Comput. Surv. **2**, 94–128.

Kramer, H. P., and Mathews, M. V. (**1956**). ''A Linear Coding for Transmitting a Set of Correlated Signals,'' IRE Trans. Inf. Theory **IT-2/3**, 41–46.

Moore, G. P. (**1992**). ''Piano trills,'' Music Percept. **9**(3), 351–359.

Moore, G. P., Hary, D., and Naill, R. (**1988**). ''Trills: Some initial observations,'' Psychomusicology **7**, 153–162.

Palmer, C. (**1996**). ''Anatomy of a performance: Sources of musical expression,'' Music Percept. **13**, 433–453.

Puckette, M. S. (**1996**). ''Pure Data,'' in *Proceedings of the International Computer Music Conference*, San Francisco, International Computer Music Association, pp. 269–272.

Sandell, G. J., and Martens, W. L. (**1995**). ''Perceptual Evaluation of Principal Component-Based Synthesis of Musical Timbres,'' J. Audio Eng. Soc. **43**, 1013–1028.

Shonle, J. I., and Horan, K. E. (**1976**). ''Trill threshold revisited,'' J. Acoust. Soc. Am. **59**, 469–471.

Smaragdis, P. (**2001**). ''Redundancy Reduction for Computational Audition, a Unifying Approach,'' Ph.D. thesis, Massachusetts Institute of Technology, Media Laboratory, Cambridge, MA.

Stapleton, J. C., and Bass, S. C. (**1988**). ''Synthesis of musical notes based on the Karhunen-Loeve transform,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-36**, 305–319.

Therrien, C. W. (**1989**). *Decision Estimation and Classification* (Wiley, New York).

Torkkola, K. (**1999**). ''Blind separation for audio signals—are we there yet?'' Proc. 1st Int. Workshop Indep. Compon. Anal. Signal Sep., Aussois, France, pp. 239–244.

Echolocation signals of dusky dolphins (*Lagenorhynchus obscurus*) in Kaikoura, New Zealand

Whitlow W. L. Au[a)]
Marine Mammal Research Program, Hawaii Institute of Marine Biology, University of Hawaii, P.O. Box 1109, Kailua, Hawaii 96734

Bernd Würsig[b)]
Marine Mammal Research Program, Texas A & M University, 4700 Avenue U, Building 303, Galveston, Texas 77551

(Received 29 August 2003; revised 23 January 2004; accepted 28 January 2004)

An array of four hydrophones arranged in a symmetrical star configuration was used to measure the echolocation signals of the dusky dolphin (*Lagenorhynchus obscurus*) near the Kaikoura Peninsula, New Zealand. Most of the echolocation signals had bi-modal frequency spectra with a low-frequency peak between 40 and 50 kHz and a high-frequency peak between 80 and 110 kHz. The low-frequency peak was dominant when the source level was low and the high frequency peak dominated when the source level was high. The center frequencies in the dusky broadband echolocation signals are among the highest of dolphins measured in the field. Peak-to-peak source levels as high as 210 dB *re* 1 μPa were measured, although the average was much lower in value. The levels of the echolocation signals are about 9–12 dB lower than for the larger white-beaked dolphin (*Lagenorhynchus albirostris*) which belongs to the same genus but is over twice as heavy as the dusky dolphins. The source level varied in amplitude approximately as a function of the one-way transmission loss for signals traveling from the animals to the array. The wave form and spectrum of the echolocation signals were similar to those of other dolphins measured in the field.
 [DOI: 10.1121/1.1690082]

PACS numbers: 43.80.Ev, 43.80.Ka, 43.80.Jz [FD] Pages: 2307–2313

I. INTRODUCTION

The dusky dolphin (*Lagenorhyncus obscurus*) is only found in the Southern hemisphere inhabiting cold-temperate coastal waters around New Zealand, western South Africa, and South America (Gaskin, 1968; Van Waerebeek *et al.*, 1995; Van Waerebeek and Würsig, 2002). Dusky dolphins are relatively small, reaching sexual maturity at about 1.75 m and rarely exceed 100 kg in weight (Van Waerebeek and Würsig, 2002). One area of New Zealand where dusky dolphins are commonly found is the waters of Kaikoura in the South Island. In this area, dusky dolphins occur in groups of generally 40–200 animals, but can reach to well over 1000 animals during April to May (Würsig *et al.*, 1997). The interaction of the subtropical convergence and the 1000-m-deep Kaikoura Canyon create a productive upwelling relatively close to shore (Robertson *et al.*, 1978), affecting the mesopelagic micronekton (small fishes, squid, and shrimp) in the diel vertical migrating deep scattering layer. Dusky dolphins forage at night on this deep scattering layer (Würsig *et al.*, 1989) as it migrates vertically from deep waters (>500 m) to within 29–49 m of the surface between 2300 and 0100 h local time (Benoit-Bird *et al.*, 2004).

Echolocation no doubt plays an important role in dusky dolphin natural history, especially when foraging at night for relatively small prey over the Kaikoura submarine canyon. Dusky dolphins in other parts of the world (Würsig and Würsig, 1980), including some other parts of New Zealand (Markowitz *et al.*, in press) also prey on schooling fish during daylight hours, no doubt utilizing their echolocation capabilities. However, there is very little information on the structure and form of echolocation signals utilized by dusky dolphins, and therefore a description of such signals near the Kaikoura Peninsula is the subject of this paper.

The echolocation characteristics of delphinid species have been studied primarily in captivity (Au, 1993). Measurements from stationary dolphins in captivity have shown that echolocation clicks are emitted in a directional beam and signals measured off-axis are distorted with respect to the signals measured along the major axis of the beam. Therefore, it is very difficult to obtain accurate measurements of free-ranging, fast moving dolphins in the wild. Another complicating factor is associated with the broadband nature of echolocation signals and the possibility that the center frequency of echolocation clicks may vary with the intensity of the emitted clicks. Au *et al.* (1985) found that higher intensity clicks emitted by a beluga whale (*Delphinapterus leucas*) in Hawaii had higher frequencies than the lower intensity clicks used by the same animal in San Diego Bay. Au *et al.* (1995) also found a nearly linear relationship between the center frequency and source level of a false killer whale (*Pseudorca crassidens*), i.e., the higher the source level, the higher the center frequency of the emitted clicks. Therefore, any measurements of the spectra of echolocation clicks should be accompanied by an estimate of the source levels.

[a)]Electronic mail: wau@hawaii.edu
[b)]Electronic mail: wursigb@tamug.edu

II. PROCEDURE

A. Measurement system

A four-hydrophone array with the hydrophones arranged as a symmetrical star was used to measure the echolocation signals of the dusky dolphin. Such a sensor geometry was used successfully by Aubauer (1995) in tracking echolocating bats in the field and by Au and various colleagues (Au and Benoit-Bird, 2003; Au and Herzing, 2003; Rasmussen *et al.*, 2002). The array structure resembled the letter ''Y,'' with each arm being 45.7 cm in length and separated by an angle of 120°. The arms of the array were constructed out of 1.27-cm o.d. PVC pipe with a spherical hydrophone connected to the end of each pipe and the cable running through the center of the pipe. Another hydrophone was connected at the geometric center of the ''Y.'' The PVC pipes fit into holes drilled into a 2.54-cm-thick delrin plate. The range of a sound source can be determined by measuring the time of arrival differences between the signal at the center and the three other hydrophones. A detailed description of the array and the location technique can be found in the appendix of Au and Herzing (2003).

An underwater housing connected to the back of the hydrophone mounting plate contained an amplifier and line-driver for each of the hydrophones. A CCD video camera in an underwater housing was mounted next to the center hydrophone. A multi-conductor cable, 7 m in length, consisting of five coaxial lines and two dc power lines, connected the array to an adjustable amplifier-filter box containing a power supply.

Echolocation signals were digitized with two Gage-1210, 12-bit dual simultaneous sampling data acquisition boards that were connected to a ''lunch box'' computer via two EISA slots. The data acquisition system operated at a sample rate of 500 kHz with a pretrigger capability. When the computer signaled the Gage-1210 to collect data, four channels of acoustic signals were simultaneously and continuously digitized, with the results going into separate circular memories on each Gage-1210 board. When an echolocation signal was detected by the center hydrophone, it triggered the data acquisition board. Two hundred pretriggered points and 200 post trigger points were collected for each channel and downloaded into the computer. A maximum of 80 clicks could be downloaded for each episode before the data had to be stored on the hard drive. A specially constructed ISA board was also used to measure the time interval between the clicks being acquired and to cause a light emitting diode to flash indicating that clicks were being captured. The interclick interval data were also downloaded and stored on the hard drive. The time of capture (to the closest 18-ms interval of the computer timing system) of each click was also saved and stored on the hard drive.

Acoustic measurements of echolocation signals were conducted from an 8.8-m motorized catamaran. All the electronics including the lunch-box computer were housed on the deck of the catamaran. The boat was driven to the vicinity of a dolphin pod and then the engine was turned off, leaving the boat to drift. The four-hydrophone array was deployed by attaching a wooden rod (2.54-cm diam) to it and the array assembly was held over the side of the boat. The center of the array was lowered to a depth of 1–1.5 m below the surface. Dusky dolphins often milled about 20–30 m from the boat, and individuals would frequently dash towards the array, continuously echolocating as they did so.

III. RESULTS

A total of 23 files containing 1456 echolocation clicks were collected representing 23 instances in which a dolphin made a run at the hydrophone array. The average number of clicks collected per file was slightly over 63. Only echolocation events in which the amplitude of the signals received by the center hydrophone was either the highest or within 3 dB of the highest were accepted for analysis. This criterion was chosen to ensure that a dolphin beam was directed at the array. The beam patterns measured for three different odontocete species (Au, 1993; Au *et al.*, 1995) indicate that when the major axis of the beam is directed to within ±5° of the center hydrophone, the signal received by the center hydrophone will either have the highest or will be within 3 dB of the highest amplitude for echolocation signals for ranges between approximately 5 and 10 m between the array and dolphins. After applying the 3-dB criterion, another level of inspection was done. Each signal passing the 3-dB criterion but was outside the 5–10-m range was examined visually for any distortions. Results from the bottlenose dolphin, *Tursiops truncatus*, the beluga whale (Au, 1993), and the false killer whale (Au *et al.*, 1995) indicate that off-axis click often have sudden reversals in the waveform or large asymmetry between the positive and negative excursions or longer intervals between successive cycles in the waveform when compared to on-axis clicks. These clicks that are off axis by more than about ±5° will appear distorted and can be easily distinguished from on-axis clicks. A total of 618 clicks met the appropriate criteria.

Four spectra of echolocation click trains are presented in a waterfall format in Fig. 1. The spectra suggest that portions of the signals in a click train can be relatively stable in shape but also include portions that are highly variable and complex. Most of the energy in the spectra is between 30 and 130 kHz, much higher than for spectra of echolocation signals measured for most other dolphins in the field (Au and Herzing, 2003; Rasmussen *et al.*, 2002). Smaller, nonwhistling odontocetes species such as the Hector's dolphin, *Cephalorhynchus hectori* (Dawson, 1988), harbor porpoise, *Phocoena phocoena* (Kamminga and Wiersma, 1981), Commerson's dolphin, *Cephalorhynchus commersonii* (Kamminga and Wiersma, 1981), finless porpoise, *Neophocaena phocaenoides* (Kamminga, 1988), and Dall's porpoise, *Phocoenoides dalli* (Hatakeyama and Soeda, 1990) emit narrowband echolocation signals that have slightly higher peak and center frequencies than the dusky dolphin.

Most of the clicks emitted by dusky dolphins had bimodal frequency spectra. The secondary peaks vary in shape from a slight ''bump'' to a clearly defined peak. A secondary peak existed if the slope of the spectrum for two consecutive increasing frequency bins changed from positive to zero or negative values. A bimodal spectrum is defined as one in which the amplitude of the secondary peak is greater than

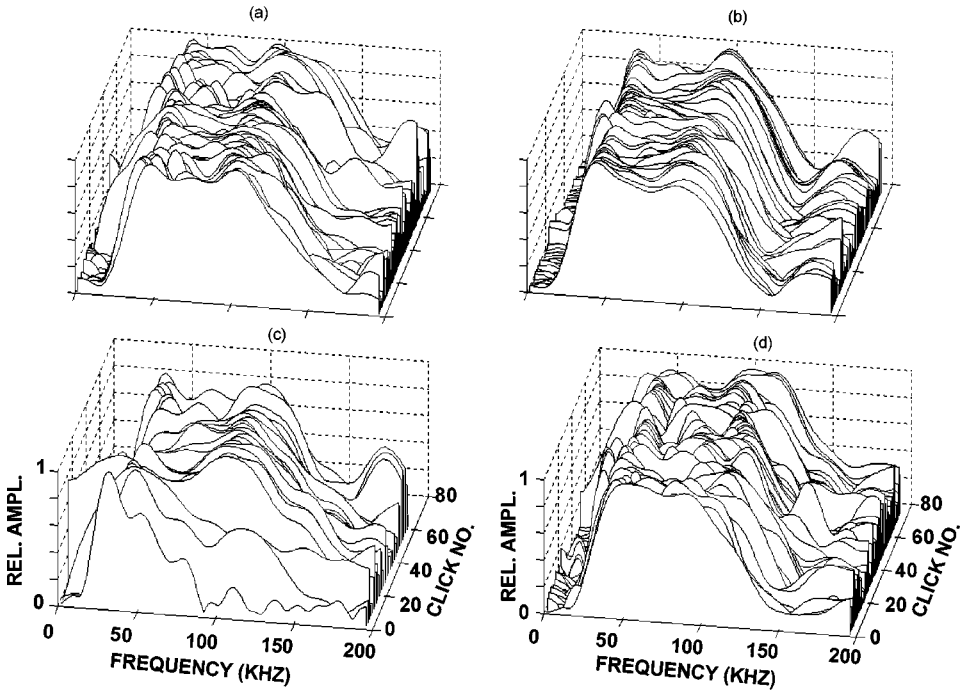


FIG. 1. Four waterfall spectra of dusky dolphin echolocation click trains.

$\frac{1}{2}$ the amplitude of the primary peak. Au *et al.* (1995) separated echolocation signals into four types based on the number of peaks in the frequency spectra. Type 1 and 2 signals had low-peak frequencies (below 64 kHz), type 1 being unimodal and type 2 being bimodal. Type 3 and 4 signals had high-peak frequencies (above 64 kHz), type 3 being bimodal and type 4 being unimodal.

Examples of representative signals of types 1–4 are shown in Fig. 2 with the waveforms on the left and the frequency spectra on the right. All the clicks were very brief, generally less than 70 μs in duration, with broad frequency spectra. The percentage of the various type signals are also shown next to the waveforms. Clicks with bimodal spectra are obvious in the spectra plots; most of the clicks had bimodal spectra (88.3%). The type 1 signals occurred 8.9% of the time and the type 4 signals occurred only 2.9% of the time. Some of the bimodal spectra have relatively high-peak frequencies (>90 kHz) whereas some have low peak frequencies (<40 kHz). The click waveforms resemble those of other odontocetes such as bottlenose dolphins, beluga whales (Au, 1993), false killer whales (Au *et al.*, 1995), and white-beaked dolphins (Rasmussen *et al.*, 2002).

The peak-to-peak source level as a function of range between an echolocating dolphin and the array is shown in Fig. 3. As the dolphin's range to the array decreased, the source level also decreased. The solid curve in Fig. 3 is a regression curve represented by the equation

$$\mathrm{SL} = 177.8 + 20\,\mathrm{Log}(\mathrm{R}), \tag{1}$$

where SL is the source level in dB *re* 1 μPa and R is the range in meters. The increase in source level as a function of 20 Log R compensates for the one-way spherical spreading loss as the signal propagates outward from a transmitter. The highest amplitude echolocation signal was 210 dB emitted by a dusky dolphin at a range of 25 m. The results also suggest that the dolphins were echolocating on the hydrophone array and not on some other objects since the source level decreased as the range to the array decreased. Also the interclick intervals were always greater than the two-way travel time from the animals to the array and back, which is consistent with the notion that the dolphins were echolocat-

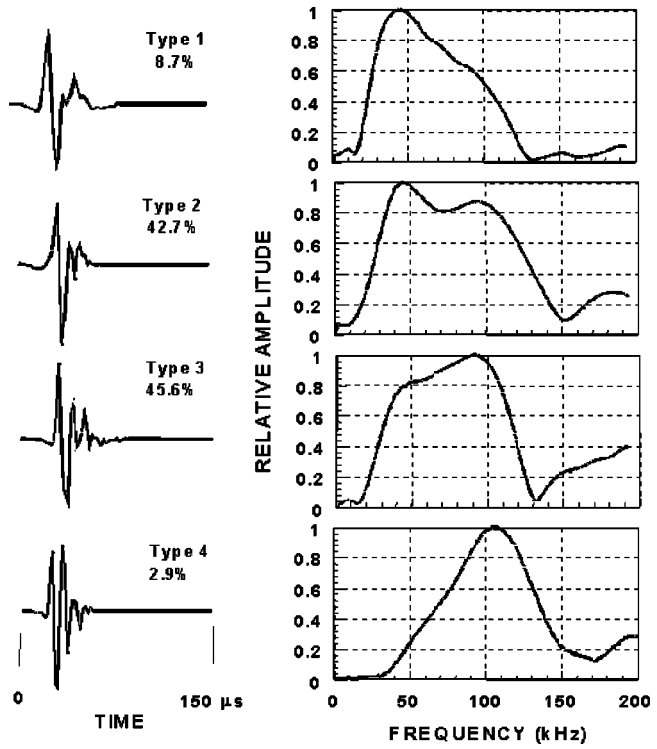


FIG. 2. Examples of some representative echolocation signal waveforms and spectra emitted by the dusky dolphin in Kaikoura. Type 1 and 2 signals have low peak frequencies (<64 kHz) with type 1 being unimodal and type 2 being bimodal. Types 3 and 4 have high peak frequencies (>64 kHz) with type 3 being bimodal and type 4 being unimodal. A bimodal signal is one in which amplitude of the secondary peak is greater than $\frac{1}{2}$ the amplitude at the peak frequency.

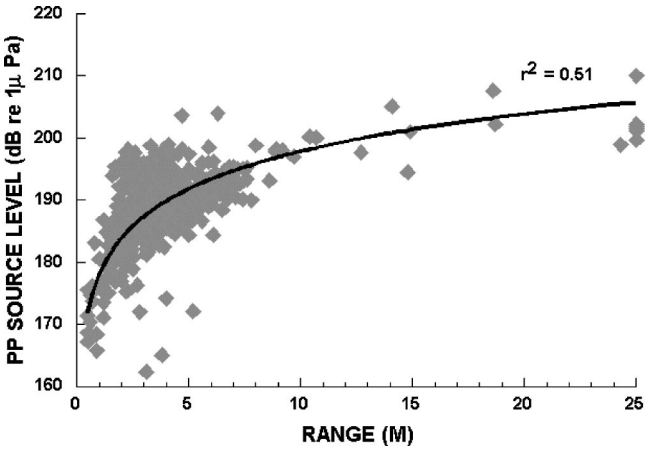


FIG. 3. Scatter plot of source level as a function of the range between an echolocating dolphin and the hydrophone array. The solid curve represents the one-way spherical spreading curve-fitted to the data in a least-square with $r^2=0.74$.

ing on the array. The fitted curve was constrained to vary as 20 Log R; however, the ''best-fit'' logarithm's curve would be very similar to the 20 Log R fit, with the r^2 values differing only in the third decimal place.

The distributions of peak and center frequencies of the echolocation signals are shown in Fig. 4. Peak frequency is defined as the frequency at which the frequency spectrum of a signal has its maximum amplitude. Center frequency is defined as that frequency which divides the energy in a frequency spectrum into two equal parts (the centroid of the spectrum). The peak frequency histogram has a low-frequency peak between 50 and 60 kHz and a high-frequency peak an octave higher between 100 and 110 kHz, reflecting the bimodal characteristics of the dusky echolocation signals. The mean and standard deviation of the peak frequency were 73.8 ± 27.3 kHz ($n=618$).

The center frequency histogram has a well-defined peak between 90 and 100 kHz. Ninety-two percent of the signals had center frequencies greater than 80 kHz and 66% had center frequencies greater than 90 kHz. The center frequency extends to much higher frequencies than the peak frequency and this property is indicative of signals with bimodal spectra; the secondary peak causes the spectrum to be broader, introducing a large asymmetry in the spectrum. Center frequency is a more representative measure of signals with bimodal spectra (Au *et al.*, 1995), since a slight shift in the spectrum could move the peak frequency over an octave away. The mean and standard deviation of the center frequency were 80.5 ± 8.7 kHz ($n=618$).

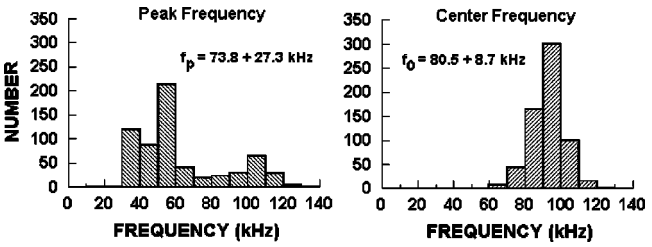


FIG. 4. Histogram of peak and center frequencies of dusky dolphin echolocation signals.

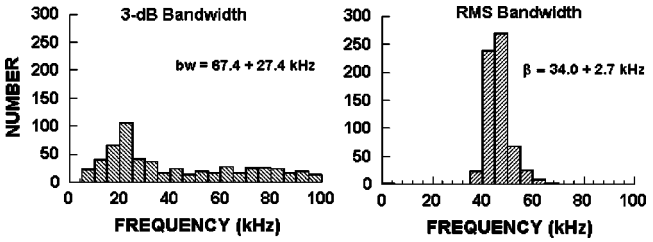


FIG. 5. Histogram of 3-dB and rms bandwidths of the dusky dolphin echolocation signals.

Histograms of the 3-dB bandwidth and the rms bandwidth are shown in Fig. 5. The 3-dB bandwidth is the width of the frequency band between the two points that are 3 dB lower than the maximum amplitude of a spectrum. The 3-dB points are also referred to as the half-power points since a level 3 dB below the maximum in a continuous signal is exactly half of the power or energy in a signal. Its distribution is rather scattered with a peak between 20 and 25 kHz, but with distributions reaching out to 100 kHz. The mean and standard deviation of the 3-dB bandwidth were 67.4 ±27.4 kHz ($n=618$).

The rms bandwidth is the standard deviation of the spectrum about the center frequency and its distribution is clustered between 40 and 50 kHz with 80% of the signals having bandwidth in this range. The 3-dB bandwidth for bimodal spectra can often provide a misrepresentation of the width of the signal since the bandwidth might cover only the frequency range about the peak frequency. The rms bandwidth is probably a better measure of the width of signals with bimodal spectra since the effects of local maxima and minima are accounted for in the calculation. The mean and standard deviation of the rms bandwidth were 34.0 ± 8.7 kHz ($n=618$).

An analysis examining the relationship between center frequency and peak-to-peak source levels indicated that there is not a direct relationship between center frequency and source level. However, when examining the average source levels for the different signal types (Fig. 6) there seems to be a trend showing that the higher frequency signals tend to have higher source levels. By the manner in which the different types of signals were defined, the type 1 signal will

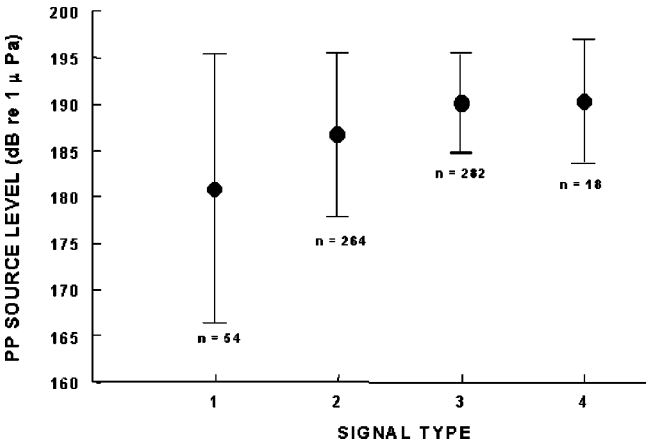


FIG. 6. The mean and standard deviation of source levels for the different signal types.

TABLE I. Results of the statistical test of the data presented in Fig. 6.

(a) One-way ANOVA

Source	DF	SS	MS	F	P
Corrected model	3	4484.93	1495.00	21.77	0.00
Error	614	42165.80	68.70		
Corrected total	46650.77				

(b) Dunnett C multiple comparison

Variable 1	Variable 1		Mean difference	Standard error	95% Confidence Interval	
					L. bound	U. bound
Type 1	Type 2	$p<0.05$	−5.84	1.24	−11.28	−0.39
	Type 3	$p<0.05$	−9.21	1.23	−14.52	−3.90
	Type 4	$p<0.05$	−9.40	2.26	−16.31	−2.49
Type 2	Type 1	$p<0.05$	5.84	1.24	0.39	11.28
	Type 3	$p<0.05$	−3.38	0.71	−5.04	−1.71
	Type 4	ns	−3.56	2.02	−8.29	1.16
Type 3	Type 1	$p<0.05$	9.21	1.23	3.90	14.52
	Type 2	$p<0.05$	3.38	0.71	1.71	5.04
	Type 4	ns	−0.19	2.01	−4.76	4.38
Type 4	Type 1	$p<0.05$	9.40	2.26	−2.49	16.31
	Type 2	ns	3.56	2.02	−1.16	8.29
	Type 3	ns	0.19	2.01	−4.38	4.76

have the lowest center frequency and the center frequency will progressively increase, with type 4 having the highest frequency. The mean source level of the type 1 signals were significantly different ($p<0.05$, one-way anova) from the that of all the other signal types (Table I). The mean source level of the type 2 signals were significantly different ($p<0.05$, one-way anova) from that of the type 1 and 3 signals and the type 4 signals were significantly different ($p<0.05$) from the type 1 signals.

IV. DISCUSSION AND CONCLUSIONS

The variation of source level as a function of 20 Log R range for the echolocation signals produced by dusky dolphins is consistent with echolocation signals of other dolphins that have been measured in the field with the symmetrical star array (Au and Benoit-Bird, 2003). Au and Benoit-Bird attributed variation of the source level as a function of range to a method of obtaining a time-varying gain function in the dolphin sonar system. The strength of echoes from a fish school containing many individual targets will decrease with range as a function of 20 Log R (Urick, 1983). Therefore, the dolphin sonar system compensates for the energy lost in the backscattered signal by increasing its source level. This time-varying gain phenomenon is not specific to a particular dolphin population or any geographic region but has been confirmed by measurements in Iceland with the white-beaked dolphins, *Lagenorhynchus albirostris* (Rasmussen *et al.*, 2002), in the Bahamas with Atlantic spotted dolphins, *Stenella frontalis* (Au and Herzing, 2003), and in British Columbia with killer whales, *Orcinus orca* (Au *et al.*, 2004).

The source levels utilized by the dusky dolphin are much lower than the signals used by the white-beaked dolphin of the same genus. The difference in source levels can be determined by the value of the constant in Eq. (1), which is 177.8 dB for the dusky dolphin and 189.6 for the white-beaked dolphin (Rasmussen *et al.*, 2002). These values represent a difference of 11.8 dB, measured with the same array and electronic system for both species. Also, the maximum source level recorded for the dusky dolphin was 210 dB *re* 1 μPa compared with 219 dB for white-beaked dolphin (Rasmussen *et al.*, 2002). The difference in source levels may not be surprising when comparing the sizes of the two species. The white-beaked dolphin has a robust body with a short thick rostrum. Adult white-beaked dolphins weigh between 220 and 350 kg and have lengths between 2.2 and 2.8 m (Reeves *et al.*, 1999) compared to approximately 100 kg and 1.75 m for the dusky dolphin. Therefore, the white-beaked dolphin is over twice as heavy as the dusky dolphin, yet only about 1 m longer, but with a much wider girth. Unfortunately, information on the relationship of animal size and source levels of echolocation signals does not exist. However, the Hector's dolphin, one of the smallest dolphins, has been reported to emit echolocation signals with typical source levels of about 150 and 175 dB (Dawson, 1988). The harbor porpoise, also one of the smallest odontocetes, was observed to emit average source levels between 165 and 169 dB while echolocating a target at a range of 25 m (Kastelein *et al.*, 1999). These levels are much lower than those used by the dusky dolphin.

Dusky dolphins project broadband, short duration echolocation signals similar to those of other odontocetes that also produce whistles. Prior to this study, the symmetrical star array had been used to measure the echolocation signals of five different species of dolphins in the field. The majority of the echolocation signals had bimodal spectra. The echolocation signals of the dusky dolphins in the waters of Kaikoura also exhibited a strong bimodal tendency, with

88.3% of the signals being bimodal. It is becoming apparent that bimodality is an inherent feature in the generation mechanism of dolphins. Type 3 signals were emitted the most (45.6% of the time); however, the number of type 2 signals was close behind, occurring 42.7% of the time.

The frequency characteristics of bimodal echolocation signals are best described by their center frequency and rms bandwidth, rather than peak frequency and 3-dB or half-power bandwidth. The standard deviation of the peak frequency values of 27.3 kHz is three times as high as the standard deviation of 8.7 kHz for center frequency. The difference in standard deviation is even greater when comparing 3-dB bandwidth with the rms bandwidth. The variance in the 3-dB bandwidth is relatively high with a standard deviation of 27.4 kHz compared with a standard deviation of only 2.7 kHz for the rms bandwidth. These results suggest that more stable measures of the spectra of echolocation signals are center frequency and rms bandwidth.

Most of the signals had center frequencies between 90 and 110 kHz, which represent some of the highest center frequencies for broadband signals emitted by free-ranging delphinids. This characteristics of high center frequency may be related to the size of the dusky dolphins. The general rule of thumb in sound production is the smaller the animal the higher the frequency of sounds produced. Wang *et al.* (1995) found a strong correlation between the maximum frequency of whistles and the body lengths of seven species of dolphins.

The broad bandwidth of the echolocation signal provides a good range resolution capability (Au, 1993) that should enable dusky dolphins, when foraging for small schooling fish (Würsig and Würsig, 1980), to perform fine target discrimination in a shallow water environment where bottom reverberation can be troublesome. Good range or time resolution would also be beneficial for dusky dolphins when foraging on mesopelagic prey such as myctophid fish and small squid (Cipriano, 1992). Although there does not seem to be a relationship between source level and center frequency as for the false killer whale (Au *et al.*, 1995), Au and Herzing (2003) found that only when the source level increases beyond 210 dB re 1 μPa did a relationship between center frequency and source level emerge. The highest level signal for the dusky dolphin was only 210 dB. Perhaps if higher source levels were used, a relationship between center frequency and source level may emerge. However, there is a definite trend in which the source level increased as the signal type increased from 1 to 3, which seems to indicate that there should be some kind of relationship between center frequency and source level. Grouping the signals into types definitely helped in bringing out the relationship between source levels and center frequency.

The results of this study also clearly demonstrate the utility of using a multi-hydrophone array to measure echolocation signals of dolphins in the wild. The symmetrical star array used in this study is relatively compact and easy to handle, and can provide information on whether a specific received echolocation signal originated in the vicinity of the major axis of the animal's transmission beam. Time of arrival differences between hydrophones were easily ascertained because of the rapid onset of the echolocation signals. Our results have demonstrated the value of using an array to obtain reliable data on echolocation signals in the field.

ACKNOWLEDGMENTS

For support in the field, we thank Ian Bradshaw, Jackie Wadsworth, and Lynnette and Dennis Buurman of ''Dolphin Encounter'' Kaikoura, New Zealand. The first author thanks Kelly Benoit-Bird for her encouragement to add a ecological perspective to this ms. He would also like to thank Kelly and Kimberly Andrews for their help in the Anova statistics. Thanks also to author Dr. Roland Aubauer for his suggestion of using a symmetrical star geometry and for various discussions associated with his research on detection of flying bats. The assistance of Michiel Schotten in testing and calibrating the array is also appreciated. For ongoing field and laboratory logistics, we thank the Edward Percival Field Station of the University of Canterbury, Jack vanBerkel, station manager. This work was partially funded by the Office of Naval Research, Dr. Robert Gisiner, program manager; the Center for Field Studies of Earthwatch; the Marlborough District Council; the New Zealand Department of Conservation; and a Fulbright research/teaching award to the second author. This is HIMB Contribution No. 1177.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer-Verlag, New York).

Au, W. W. L., and Benoit-Bird, K. J. (**2003**). ''Automatic gain control in the echolocation system of dolphins,'' Nature (London) **423**, 861–863.

Au, W. W. L., and Herzing, D. L. (**2003**). ''Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*),'' J. Acoust. Soc. Am. **113**, 598–604.

Au, W. W. L., Carder, D. A., Penner, R. H., and Scronce, B. L. (**1985**). ''Demonstration of Adaptation in Beluga whale Echolocation Signals,'' J. Acoust. Soc. Am. **77**, 726–730.

Au, W. W. L., Ford, J. K. B., Horne, J. K., and Allman, K. A. (**2004**). ''Echolocation signals of free-ranging killer whales (Orcinus orca) and modeling of foraging for chinook salmon (Oncorhynchus tshawytscha),'' J. Acoust. Soc. Am. **115**, 901–909.

Au, W. W. L., Pawloski, J. L., Nachtigall, P. E., Blonz, M., and Gisner, R. C. (**1995**). ''Echolocation signals and transmission beam pattern of a false killer whale (*Pseudorca crassidens*),'' J. Acoust. Soc. Am. **98**, 51–59.

Aubauer, R. (**1995**). ''Korrelationsverfahren zur Flugbahnverfolgung echoortender Fledemause,'' unpublished Dipl.-Ing, Technishen Hochschule Darmstadt, Darmstadt.

Benoit-Bird, K. J., Wursig, B., and FcFadden, C. J. (**2004**). ''Dusky dolphin (*Lagenorhynchus obscurus*) foraging in two different habitats: Active acoustic detection of dolphins and their prey,'' in press (April), Mar. Mamm. Sci.

Cipriano, F. (**1992**). ''Behavior and occurrence patterns, feeding ecology, and life history of dusky dolphins (*Lagenorhynchus obscurus*) off Kaikoura, New Zealand,'' unpublished Ph.D. dissertation, University of Tucson, Tucson, AZ.

Dawson, S. M. (**1988**). ''The high frequency sounds of free-ranging Hector's dolphin, *Cephalorhynchus hectori*,'' Rep. Int. Whal. Commission, Special Issue **9**, 339–341.

Gaskin, D. E. (**1968**). ''Distribution of Delphinidae (Cetacea) in relation to sea surface temperatures off Eastern and Southern New Zealand,'' N. Z. J. Mar. Freshwater Res. **2**, 527–534.

Hatakeyama, Y., and Soeda, H. (**1990**). ''Observation of porpoises taken in salmon gillnet fisheries,'' in *Sensory Abilities of Cetaceans*, edited by J. A. Thomas and R. Kastelein (Plenum, New York), pp. 269–281.

Kamminga, C. (**1988**). ''Echolocation signal types of odontocetes,'' in *Animal Sonar: Processes and Performance*, edited by P. E. Nachtigall and P. W. B. Moore (Plenum, New York), pp. 9–22.

Kamminga, C., and Wiersma, H. (**1981**). ''Investigations of cetacean sonar II. Acoustical similarities and differences in odontocete sonar signals,'' Aquat. Mamm. **8**, 41–62.

Kastelein, R. A., Au, W. W. L., Rippe, H. T., and Schooneman, N. M. (**1999**). ''Target detection by an cholocating harbor porpoise (*Phocoena phocoena*),'' J. Acoust. Soc. Am. **105**, 2493–2498.

Markowitz, T. M., Harlin, A. D., Wursig, B., and J, M. C. (**in press**). ''Dusky dolphin foraging habitat: Overlap with aquaculture in New Zealand,'' Aquatic Conservation.

Rasmussen, M. H., Miller, L. A., and Au, W. W. L. (**2002**). ''Source levels of clicks from free-ranging white beaked dolphins (*Lagenorhynchus albirostris* Gray 1846) recorded in Icelandic waters,'' J. Acoust. Soc. Am. **111**, 1122–1125.

Reeves, R. R., Smeenk, C., Kinze, C. C., Brownell, R. L., and Lien, J. (**1999**). ''White-beaked dolphin, Lagenorhynchus albirostris Gray, 1846,'' in *Handbook of Marine Mammals Vol 6* S. H.

Robertson, D. A., Roberts, P. E., and Wilson, J. B. (**1978**). ''Mesopelagic faunal transition across the Subtropical Convergence east of New Zealand,'' N. Z. J. Mar. Freshwater Res. **12**, 295–312.

Urick, R. J. (**1983**). *Principles of Underwater Sound* (McGraw–Hill, New York).

Van Waerebeek, K., Van Bree, P. J., and Best, P. B. (**1995**). ''On the identity of Prodelphinus petersii (Lütkin, 1889) and records of dusky dolphin Lagenorhynchus obscurus (Gray, 1928) from the southern Indian and Atlantic Oceans,'' S. Afr. J. Mar. Sci. **16**, 25–35.

Van Waerebeek, K., and Würsig, B. (**2002**). ''Pacific white-sided dolphin and dusky dolphin,'' in *Encyclopedia of Marine Mammals*, edited by W. F. Perrin, B. Wursig, and J. G. M. Thewissen (Academic, San Diego), pp. 859–861.

Wang, D., Wursig, B., and Evans, W. E. (**1995**). ''Comparison of whistles among seven odontocete species,'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil, Woeden, The Netherlands), pp. 299–323.

Würsig, B., and Würsig, M. (**1980**). ''Behavior and ecology of the dusky dolphin, Lagenorhynchus obscurus, in the south Atlantic,'' Fish. Bull. **77**, 871–890.

Würsig, B., Würsig, M., and Cipriano, F. (**1989**). ''Dolphins in different worlds,'' Oceanus **32**, 71–75.

Würsig, M., Cipriano, F., Slooten, E., Constantine, R., Barr, K., and Yin, S. (**1997**). ''Dusky dolphin (*Lagenorhynchus obscurus*) off New Zealand: Status of present knowledge,'' Rep. Int. Whaling Commission **77**, 871–890.

Three-dimensional simulations of ultrasonic axial transmission velocity measurement on cortical bone models

Emmanuel Bossy,[a] Maryline Talmant, and Pascal Laugier
Laboratoire d'Imagerie Paramétrique, UMR CNRS 7623, Université Paris VI, 75006 Paris, France

(Received 24 July 2003; revised 22 January 2004; accepted 1 February 2004)

The ultrasonic axial transmission technique, used to assess cortical shells of long bones, is investigated using numerical simulations based on a three-dimensional (3D) finite difference code. We focus our interest on the effects of 3D cortical bone geometry (curvature, cortical thickness), anisotropy, and microporosity on speed of sound (SOS) measurements for different frequencies in the MHz range. We first show that SOS values measured on tubular cortical shells are identical to those measured on cortical plates of equal thickness. Anisotropy of cortical bone is then shown to have a major impact on SOS measurement as a function of cortical thickness. The range of SOS values measured on anisotropic bone is half the range found when bone is considered isotropic. Dependence of thickness occurs for cortical shell thinner than $0.5\times\lambda_{bone}$ in anisotropic bone (λ_{bone}: wavelength in bone), whereas it occurs for cortical shell thinner than λ_{bone} when anisotropy is neglected. Sensitivity of SOS along the bone axis to intracortical microporosity is shown to be approximately $-20\ \mathrm{m\,s^{-1}}$ per percent of porosity. Using homogenized porous bone, we finally show that the cortical depth that contributes to lateral wave SOS measurement is approximately 1–1.5 mm for frequencies ranging from 500 kHz to 2 MHz under classical *in vivo* measurement conditions.
 [DOI: 10.1121/1.1689960]

PACS numbers: 43.80.Ev, 43.80.Vj, 43.80.Jz, 43.80.Qf [FD] Pages: 2314–2324

I. INTRODUCTION

In recent years, quantitative ultrasound has played an increasing role in the assessment of bone status in osteoporosis. The axial transmission technique has been developed to assess cortical bone properties at skeletal sites such as the radius or tibia.[1–3] With this technique, both the emitter(s) and the receiver(s) are in contact with the skin, placed on a same side of the investigated skeletal site. A low frequency ultrasound wave of typically 250 kHz–2 MHz frequency propagates along the long axis of the cortical shell of the measured skeletal site, and the first arriving signal detected at the receiver(s) is used to derive a speed of sound (SOS) value. Such a technique has been developed with the aim of assessing cortical bone status under various pathological conditions for elderly people (osteoporosis[4–6]), children (juvenile arthritis,[7] lymphoblastic leukemia[8]), and neonates.[9] However, the primary interest has focused on osteoporosis. In osteoporosis, modifications mainly occur in the endosteal (i.e., the inner part) region of the cortex, where resorption leads to an increased porosity, trabecularization, and reduced cortical thickness and results in an increased bone fragility.[10–12] One device is currently commercially available, which enables ultrasound evaluation of the distal one-third radius and potentially of other skeletal sites including the ulna, finger phalanxes, metacarpal, or metatarsus. Several clinical studies demonstrated adequate precision for investigating skeletal status and significant fracture risk discrimination.[1,4,6]

Although both experimental[13–17] and simulations[18,19] studies have reported correlations between axial transmission SOS measurements and bone properties such as bone mineral density, cortical thickness, and elasticity modulus, the exact determinants of SOS values remain unclear. In a previous study,[18] we showed by means of two-dimensional (2D) simulations on models of isotropic homogeneous cortical bone plates that SOS values were related to cortical thickness for appropriate frequency ranges, in agreement with experimental studies.[15,16] Nevertheless, several questions still have to be answered: what is the influence of the complex 3D geometry of cortical bone on SOS measurements? Does anisotropy play a significant role in the involved propagation phenomena? What is the expected sensitivity of this technique that uses emission and detection of ultrasound on periosteal regions of bone (i.e., the external part of the cortex) to porosity increase that occurs primarily in the endosteal region of the cortex? The present study addresses these issues by means of three-dimensional (3D) simulations, taking into account for the first time the 3D tubular geometry of bone, its elastic anisotropy, and its porosity. As background, we briefly introduce axial transmission measurement principles, and recall the results issued from our previous 2D simulation study on cortical bone plates. The finite-difference method that we used to compute numerical results is then presented. We first compare our previous 2D results with results obtained using a more realistic 3D tubular geometry. We then take into account the effect of anisotropy on axial transmission measurements, that have been neglected so far in previous modeling.[16,18,19] We finally address the question of the sensitivity of the axial transmission technique to increasing cortical porosity, for ultrasonic frequencies ranging from 500 kHz to 2 MHz.

[a]Electronic mail: emmanuel.bossy@lip.bhdc.jussieu.fr

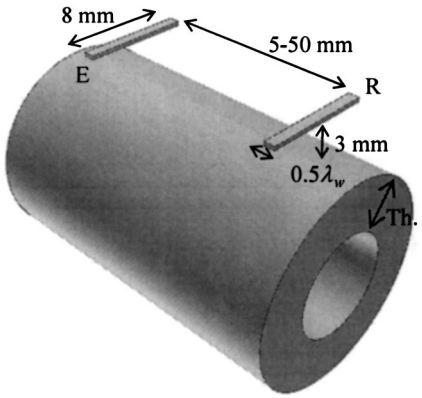


FIG. 1. Typical three-dimensional measurement configuration involved in axial transmission on cortical bone, where λ_w is the wavelength in water (or soft tissue). E: emitter; R: receiver(s). Th: cortical thickness.

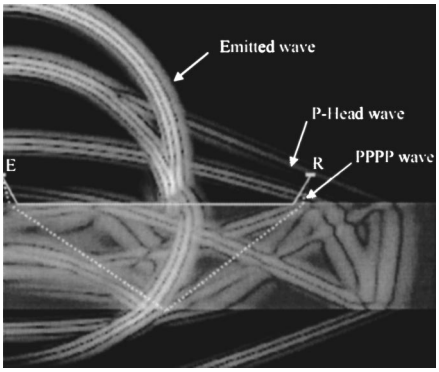


FIG. 2. Snapshot of 2D wave propagation on a cortical bone plate. Propagation paths from emitter (E) to receiver (R) are indicated for the lateral wave (or P-head wave, full line) and the PPPP wave (dotted line).

II. BACKGROUND

A. Axial transmission

Unlike transverse transmission techniques, which involve emission and reception of ultrasonic waves on opposites sides of a skeletal site, the so-called axial transmission has been developed to study propagation along the cortical shell of long bones.[1–3] In axial transmission techniques, both emitters and receivers are placed in contact with the skin on the same side of the skeletal site. Different transducer arrangements have been devised.[3,6,20] Typically, frequencies from a few hundred kHz up to 2 MHz have been used at different cortical bone sites including the radius, tibia, ulna, and phalanx.[2,3,5] *In vivo*, emitter(s)–receiver(s) distances typically range from 5 to 50 mm.[2,3,5,19] In this work, the emitter(s) and receiver(s) spatial arrangement was simulated to follow a probe design developed in our laboratory, based on an multielement array technology recently described by our group.[20] The 3D experimental configuration and the location of the transducers relatively to bone are shown on Fig. 1. The distance between bone surface and transducers is set to 3 mm throughout our study. Each individual element, either emitter or receiver, consists in a $\lambda_w/2$ long element, where λ_w is the ultrasonic wavelength in soft tissue, close to the wavelength in water. In the transverse direction, each element is 8 mm wide. Such dimensions ensure no directivity for both emitters and receivers in the longitudinal plane.

For sake of clarity, the type of propagation involved in axial transmission is now briefly recalled on a 2D isotropic plate model (Fig. 2). Further detail on the 2D model can be found in previous reports.[18,21] The quasicircular wave front emitted by a transducer placed on top of the overlying soft tissue (or an acoustically equivalent fluid) is partially transmitted into the bone, with strong refraction due to high velocity contrast between soft tissue (bulk compressional velocity $\sim 1500\ \mathrm{m\,s^{-1}}$) and bone (bulk compressional velocity $\sim 4000\ \mathrm{m\,s^{-1}}$). Complicated phenomena may occur depending on the wavelength-to-plate thickness ratio, but consistently, part of the ultrasonic energy guided along the plate is radiated back into the soft tissue, as shown on Fig. 2. In commercialized devices,[1,3] the first detectable wave radiated from bone is used to derive the SOS of the wave guided along the cortical surface. The following paragraph recalls previous results obtained from 2D simulations on isotropic elastic plates.[18]

B. Thickness effect: Results on 2D cortical bone plates

In a previous study,[18] we simulated the effect of cortical thickness on SOS axial transmission measurements, based on a 2D isotropic cortical bone plate model. We showed that for plate thicknesses larger than the compressional wavelength in bone, the first arriving signal corresponds to the radiation of a compressional bulk wave guided by the soft tissue/bone interface. This radiated wave is called lateral wave, or P-head wave. For thin plates, typically less than a quarter of a compressional wavelength, the first arriving signal corresponds to the radiation of the S_0 Lamb mode, guided by the whole cortical thickness. For intermediate thickness, the detected signal results from a complex pattern of interference between different waves. In particular, when the cortical thickness is about one compressional wavelength, the first arriving signal may be interpreted as an interference between the lateral wave and a compressional wave reflected from the bottom of the plate (Fig. 2), commonly noted PPPP wave in geophysics. For intermediate thicknesses, the shape of the first arriving signal is highly dependent on the receiver position, with a relatively weak amplitude compared to smaller and larger thicknesses. For large and thin plates, the first arriving signal corresponds to the radiation of a nondispersive wave, and phase velocity (equal to group velocity) can be accurately measured by detection of signal extrema or using a cross-correlation technique.[18] These results were in agreement with published experimental[15,16,17,19] and simulation results.[19] Nevertheless, a more realistic model must include 3D cortical geometry, along with anisotropic properties of bone material to provide a comprehensive understanding of the involved propagation characteristics.

C. Lateral wave velocity as a function of emitter–receiver distance

The velocity of the lateral wave predicted from the theory in the far field[22] corresponds to the bulk compressional velocity in the solid material. In our previous 2D

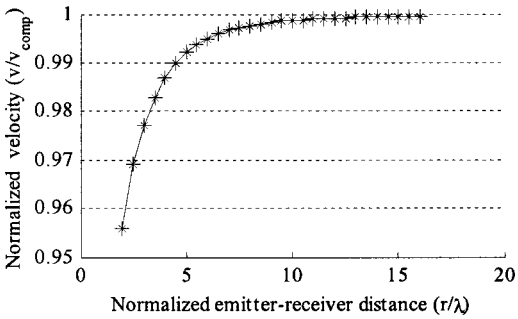


FIG. 3. Local lateral wave velocity, normalized by bulk wave compressional velocity, as a function of emitter–receiver distance to wavelength ratio.

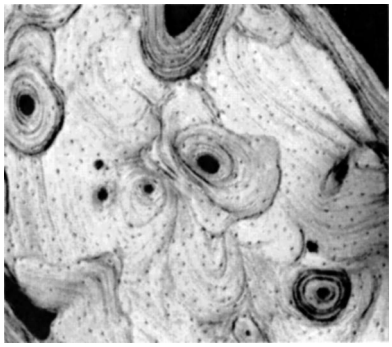


FIG. 4. Cortical bone image (1 mm^2) obtained from 400 MHz acoustic microscopy.

study,[18] we reported a near-field effect that resulted in a lateral wave velocity slightly but significantly lower. The simulations have been extended since then to 500 kHz and 2 MHz on several test materials (perspex, aluminum), which confirmed this near-field effect. This near-field effect is summarized by the single curve in Fig. 3, which plots the measured lateral wave velocity values normalized to the material compressional bulk wave velocity, as a function of emitter–receiver distances normalized to wavelength. Though relatively low, the deviation from actual compressional velocity for emitter–receiver distances of a few wavelengths can be significant in the context of clinical measurements, in particular for the comparison of measurements obtained from different devices between which the emitter–receiver distances would differ.

III. SIMULATION METHODS

A. Algorithm

Simulations were performed using a finite-difference code developed in our laboratory, which computes a numerical solution to the 3D linear elastic wave propagation. Briefly, the algorithm is based on the spatial and temporal discretization of the following first-order equations:

$$\frac{\partial \mathrm{v}_i}{\partial t} = \frac{1}{\rho(\mathbf{r})} \times \frac{\partial \mathrm{T}_{ij}}{\partial r_j}, \tag{1}$$

$$\frac{\partial \mathrm{T}_{ij}}{\partial t} = \mathrm{C}_{ijkl}(\mathbf{r}) \times \frac{\partial \mathrm{v}_k}{\partial r_l}, \tag{2}$$

where **r** is the position vector, **v** the displacement velocity, **T** the stress tensor and **C** the stiffness tensor.

Equations (1) and (2) are written using Einstein's convention for implicit summation, with all subscripts varying from 1 to 3. This set of nine equations fully describes propagation in nonabsorbing, heterogeneous, anisotropic and elastic media. The discretization of these equations according to the Virieux scheme[23,24] was chosen for its ability to accurately model propagation for both fluids and solids, and for a convenient implementation of perfectly matched layers (PML)[25] on sides and edges of the simulation mesh, essential to efficiently avoid unphysical numerical reflections. 2D propagation was simulated using an analog restricted 2D algorithm. Appropriate spatial pitch is specified for each of the following sections. The time step was automatically deduced from the stability condition required by the Virieux scheme.[23,24]

B. Materials modeling

1. Soft tissues

In this study, soft tissue (skin, marrow, etc.) is considered to be water, with a density of 1 $g\,cm^{-3}$ and an acoustic velocity of 1500 $m\,s^{-1}$. Therefore, we neglect absorption phenomena and variations in soft tissue properties.

2. Cortical bone

a. Structure. Cortical bone, also known as compact bone, is found in the shaft of long bones such as radius or femur. It forms the outer shell around the medullar canal filled with marrow. Porosity values in cortical bone range from a few percent up to 20%.[26] Cortical bone has a hierarchical architecture with many different structures on many levels of scale.[27] The basic first level structure of cortical bone is made of osteons, which are long narrow cylinders with typical diameters ranging between 100 and 300 μm (Fig. 4), approximately 10 mm long. At the second level structure, each osteon is composed of a series of contiguous circumferential lamellae. At the center of each osteon is a central canal containing blood vessels: the Haversian canal. Cortical bone is anisotropic. Its anisotropy reflects the directionality in its microstructure and ultrastructure. Human cortical bone is generally considered as an orthotropic medium, but for our purpose, we will consider that the elasticity of bone follows a simplified form of orthotropy called transverse isotropy.[28]

b. Cortical bone as a homogeneous medium. With typical shear and compressional velocities of about 2000 and 4000 $m\,s^{-1}$ in cortical bone, ultrasonic wavelengths in the MHz frequency range are in the millimeter range, much greater than any structural heterogeneity. For the purpose of our simulations at the whole bone level, cortical bone was therefore considered to be a solid material with effective stiffness at the scale of the wavelength. In a first approach (Sec. IV), anisotropy was neglected and isotropic elastic constants indicated in Table I were used. These constants were derived from shear and compressional velocity values found in the literature (approximately 1800 and 4000 $m\,s^{-1}$, respectively).[5,29,30] In this paper, a value of 1.85 $g\,cm^{-3}$ was

TABLE I. Values of stiffness constants (GPa) for isotropic and transversely isotropic cortical bone used in the numerical simulations.

	Isotropic	Anisotropic
$C_{11}=C_{22}$	29.6	21.5
C_{33}	29.6	29.6
$C_{12}=C_{23}=C_{31}$	17.6	11.5
$C_{44}=C_{55}$	06.0	06.0
C_{66}	06.0	05.0

consistently considered for cortical bone density. In a second approach (Sec. V), anisotropy was taken into account. Anisotropic elastic constant indicated in Table I were used to model transversely isotropic bone, according to values found experimentally.[28,30] Corresponding velocities are indicated in Table II. Note that the anisotropic constants in Table I correspond to effective elastic constants, obtained from experimental measurements in the MHz frequency range.[30] In particular, they do not reflect only elastic properties at the material level, but also the intracortical microporosity.

c. Modeling intracortical microporosity Porosity in cortical bone includes pores on several levels of scale: resorption cavities in the endosteal region with sizes of the order of a few microns, Haversian canals with an average diameter of 50–100 μm, and lacunae with an average diameter of 10–20 μm. In Sec. VI A, we study the effect of such a porosity network on the effective velocity values. To this goal, several hypotheses were made. The architecture of the porosity network in the cortical shell is mainly due to the Haversian canals (Fig. 5). It was modeled here by cylindrical marrow inclusions, embedded in an isotropic bony matrix, with their axis parallel to the longitudinal axis of bone. Periodic and quasirandom pores spatial distributions were generated with porosity values ranging from 0% to 20%. The mean porosity value was controlled by pores diameter, while the number of pores per mm^2 in transverse sections was held constant to a value of approximately 10 pores mm^{-2}, according to published results.[26]

C. Signal processing

Each ultrasonic emission consisted in a broadband pressure pulse. Three-dimensional simulations in Secs. IV and V involved a 1 MHz center frequency, while all other simulations involved 500 kHz, 1 MHz, and 2 MHz center frequencies. Signal processing involved the choice of a time criterion to detect the first arriving signal and to derive SOS values. It was based here on the detection of signal extrema, which led to accurate estimates of velocities for nondispersive signals, as detailed in our previous 2D study.[18] Briefly, several receivers are regularly distributed along the axis of the bone. Local velocity values were calculated by dividing the distance separating two adjacent receivers (typically 1 mm apart along bone surface) by the differences in time-of-flight measured at the two corresponding receivers (Secs. IV, V, VI B). Effective velocity values through porous bone (Sec. VI A) were simply estimated from measurements of the time-of-flight needed to propagate through a few wavelengths of porous bone.

TABLE II. Ultrasonic velocities (m s^{-1}) of bone tissue along principal directions in cortical bone.

	Isotropic bone		Anisotropic bone	
Direction	Comp	Shear	Comp	Shear
Axial	4000	1800	4000	1645
Transverse	4000	1800	3410	1800

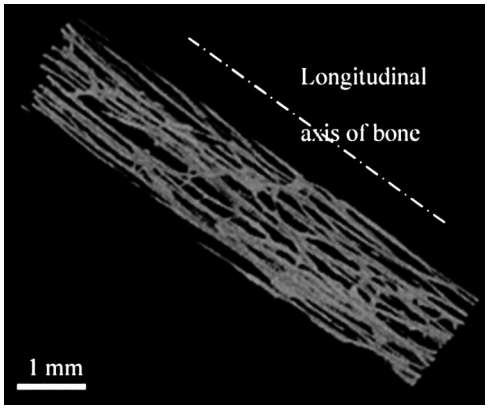


FIG. 5. Intracortical porosity architecture, obtained from 3D synchrotron microtomography (respectively, 10 μm). Only pores (mainly Haversian canals) are visualized, bone tissue having been subtracted from the image.

D. Snapshots

Each simulation snapshot represents absolute values of the displacement velocity wave field, using a logarithmic gray scale combined with transparency effects to visualize bone.

IV. 3D SIMULATIONS RESULTS: PLATE MODEL VS TUBULAR MODEL

A. Curvature effect

Previous modeling assumed 2D plates model and neglected the curvature that is encountered on real bone samples.[18,19] While curvature along the axis of bone may usually reasonably be neglected at the radius [Fig. 6(b)] or compensated for using a dedicated bidirectional transmission technique,[20] the convex curvature shown on a transverse cross section of the bone [Fig. 6(a)] may play a role in axial transmission propagation. Moreover, this curvature may be more or less pronounced depending on sites and individuals. To assess the effect of the interface curvature, we simulated SOS axial transmission measurements on an appropriate dedicated semi-infinite hemicylinder geometry, depicted in Fig. 7. This geometry has been chosen to assess the effect of the curvature of the interface under the transducers alone, independent of any potential confounding effects of the medium thickness. Simulations were performed for a 1 MHz broadband pulse, with radius of curvature r_c ranging from 2 mm up to the case of a plane interface. Such values include curvatures that may be found in the diaphysis of long bones such as radius, ulna, tibia or phalanx for instance.

In this part, cortical bone has been modeled as a homogeneous isotropic material (i.e., the porosity is neglected), with effective elastic constants given in Table I. In order to

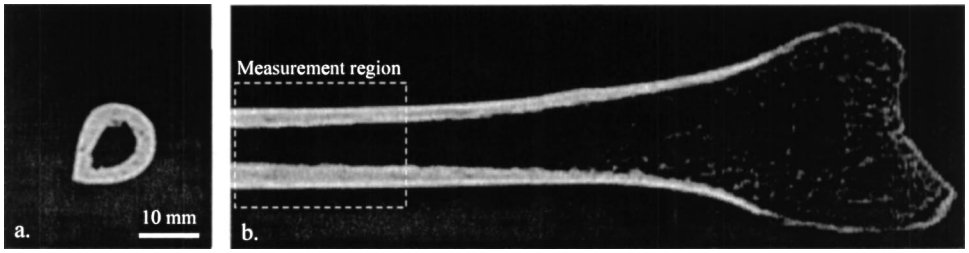


FIG. 6. Transverse (a) and longitudinal (b) x-ray computed tomography cross sections of a human radius (Siemens, Somaton 4 Plus), illustrating surface curvatures in a typical ultrasound measurement region.

ensure accurate 3D simulations, the 3D spatial pitch was set to 100 μm, corresponding to a fifteenth of the shortest wavelength involved (wavelength in water, $\lambda_w = 1.5$ mm). Typically on a 1 GHz PC, 5–10 h computational times and 500 Mbytes to 2 Gbytes RAM memory were required for each simulation. The results, expressed in terms of velocity as a function of the radius-to-wavelength ratio ($\lambda_{\text{bone(1 MHz)}} = 4$ mm), are plotted in Fig. 8, with a *y* axis range purposely set to that of Fig. 3. SOS values obtained on curved interfaces are not significantly different from those obtained on a plane interface. Although the particular 3D geometry influences most of the received signal, it has nevertheless no effect on the first part of this signal, which corresponds to propagation through the fastest path from the emitter to the receiver. This can qualitatively be observed on 2D and 3D snapshots of the wave field in Fig. 9. In particular, this tends to show that variations in SOS values like those described by others[31] or measured *in vivo* and *in vitro* in our lab (results not shown here), for different positions of the probe around the cortical circumference, cannot be attributed to differences in local curvature but are rather due to variations either in cortical thickness or bone material properties.

B. Thickness effect

The cortical thickness influence was then assessed on a 3D tubular geometry illustrated in Fig. 1. The outer radius r_{out} was set to a fixed value of 8 mm, in agreement with typical dimension found at the distal third of a human radius.

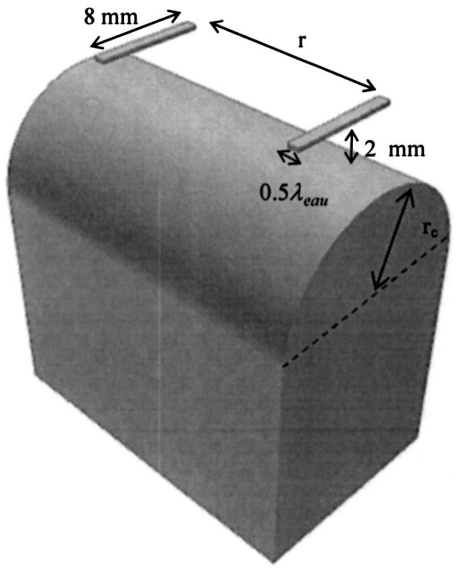


FIG. 7. 3D geometry used to assess curvature effects on axial transmission measurements.

Cortical thickness values were chosen in the 0.5–7 mm range, by changing the inner radius r_{in}. The simulation parameters were identical to those used in the previous paragraph, with similar computer requirements. Again, no significant differences were observed between 3D and 2D models, in terms of SOS of the first arriving signal. The values of velocity derived from 3D computations (see Fig. 11, isotropic case) are identical to those obtained on 2D cortical bone plates.[18] 3D and 2D snapshots [Figs. 10(a) and (b)] of the wave field indeed show strictly identical radiated patterns, when considering the first radiated wave only.

C. Discussion

For homogeneous and isotropic elastic solids, a 2D plate model is strictly equivalent to a more refined 3D tubular cylindrical model when SOS measurements of the first arriving signal only are considered. In particular, this means that the part of the tubular shell just beneath the probe is the only part of the whole cortical shell that contributes to the first arriving signal. Nevertheless, the whole signal that results from the interaction of the incident pulse and the 3D tubular geometry is, as expected, very different from the whole signal that results from the interaction of the incident pulse with the plate geometry. Measurement that would involve later contribution of the ultrasonic signal might be different on 2D and 3D geometries, depending on the measured propagation modes. Our conclusion here only applies to SOS measurements derived from the propagation of the first arriving signal on short distances (typically a few wavelength), and validates previous results issued in this context from 2D simulation studies[18,19] performed on isotropic bone mimicking models.

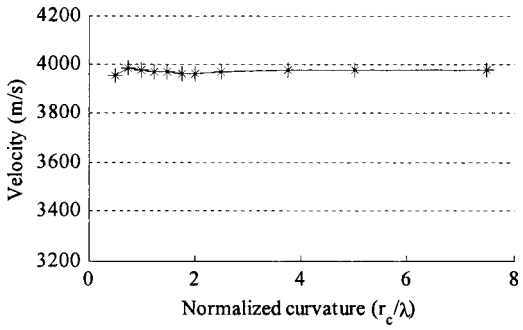


FIG. 8. Measured velocity as a function of the radius of curvature normalized to the wavelength of bulk compressional wave in bone.

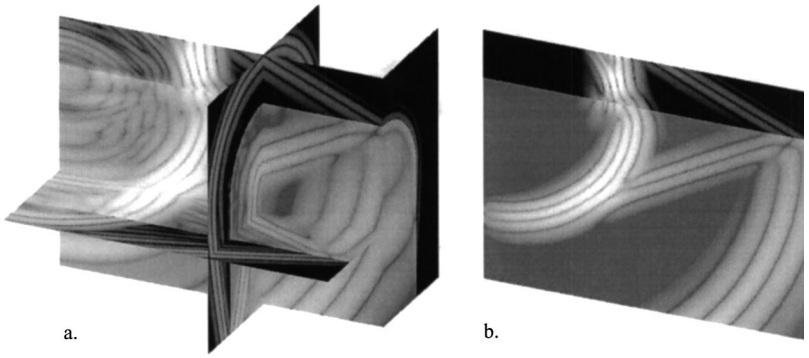


FIG. 9. 3D (a) and 2D (b) snapshots of wave propagation occurring on the geometry depicted in Fig. 6 geometry, showing identical first radiated waves. Frequency is 1 MHz, and the radius of curvature is 4 mm.

V. EFFECT OF BONE ANISOTROPY ON AXIAL TRANSMISSION MEASUREMENTS

To the best of our knowledge, cortical bone has always been modeled as an isotropic material in previous studies involving axial transmission measurement.[2,18,19] However, bone is anisotropic, as discussed in Sec. III B. Radial and tangential velocities values are approximately the same ($V_{\text{transverse}}$ typically 3400 m s^{-1}) and are significantly lower than axial velocity (V_{axial} typically 4000 m s^{-1}). To investigate the effect of anisotropy on the dependence of SOS values on cortical thickness, both 2D and 3D simulations have been performed. As expected from the results reported in Sec. IV, no difference was observed between 2D and 3D simulations, though anisotropy was taken into account. Two major differences are observed between the curves obtained in the isotropic and anisotropic case (Fig. 11).

First, the critical cortical thickness above which the SOS values is equal to the velocity of the compressional bulk wave velocity has decreased down to $\lambda_{\text{bone(axial)}}/2$ (where $\lambda_{\text{bone(axial)}}$ is the compressional wavelength along the longitudinal axis of bone), compared to a critical thickness around $\lambda_{\text{bone(axial)}}$ found in the isotropic case. As detailed in our previous 2D study,[18] the range of thickness allowing measurement of the compressional bulk wave velocity can be predicted from the occurrence of interference between the lateral wave and the PPPP wave reflected from the inner interface of the cortical shell. The difference in critical thickness observed between isotropic bone and transversely isotropic bone may be partly due to a longer propagation time required for the PPPP wave to bounce back and interfere with the lateral wave, as the radial velocity is lower than the axial velocity. Although the relative difference between axial and radial velocities expressed as $(V_{\text{axial}}-V_{\text{transverse}})/V_{\text{axial}}$ is only 15%, it has a major effect on the critical thickness which is now divided by two compared to the isotropic case. As a consequence, the sensitivity of axial transmission SOS measurements to cortical thickness occurs in a narrower thickness range, from $\lambda_{\text{bone(axial)}}/4$ to $\lambda_{\text{bone(axial)}}/2$, approximately half the range predicted for the isotropic case (Fig. 3). Given the range of cortical thickness encountered at measured skeletal site such as distal radius or tibia (cortical thickness typically ranging between 1 and 7 mm[15,17]), these results imply that the impact of cortical thickness on measured SOS values should be limited at 1 MHz or even insignificant at higher frequencies. Practically, axial transmission measurements performed at 1 MHz should be insensitive to cortical thickness larger than 2 mm, whereas measurements performed at 250 kHz should be sensitive to thickness values ranging from 4 to 8 mm. This is consistent with experimental *in vivo* studies, which showed poor or insignificant correlation between 1.25 MHz SOS measurements and cortical thickness at the radius ($r^2=0.14, p<0.05$) and tibia ($r^2=0.06$),[17] but significant higher correlation between SOS measurements at 250 kHz and cortical thickness at the tibia ($r^2=0.62, p<10^{-3}$).[15]

Second, the range of measured velocity has decreased from approximately (3200–4000 m s^{-1}) to (3600–4000 m s^{-1}). The theoretical predicted value for the nondispersive low frequency S_0 Lamb mode, in the case of propagation along a main symmetry axis (denoted x_3) of the bone material, parallel to the plate, is given by

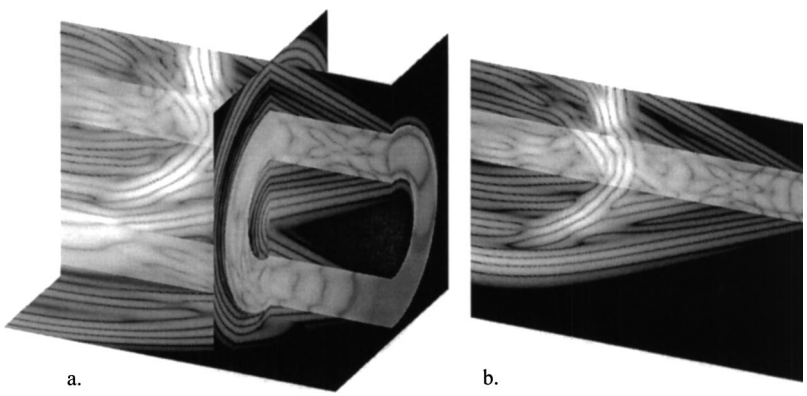


FIG. 10. 3D (a) and 2D (b) snapshots of wave propagation occurring on a tubular geometry (Fig. 1), showing identical first radiated waves as measured in axial transmission. Frequency is 1 MHz, cortical thickness is 4 mm ($r_{\text{out}}=12$ mm, $r_{\text{in}}=8$ mm).

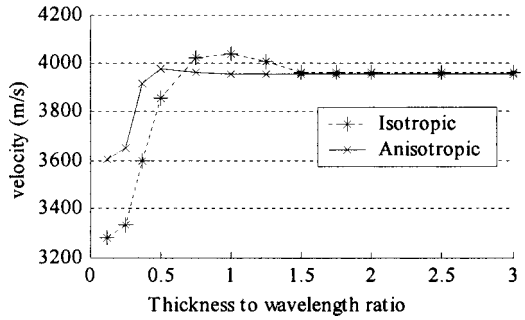


FIG. 11. Effect of anisotropy on the dependence of SOS on cortical thickness (3D computations).

$$v_{S_0} = \sqrt{\frac{c_{33}}{\rho} \times \left(1 - \frac{c_{13}^2}{c_{11} \times c_{33}}\right)}$$

$$= v_{\text{axial}} \times \sqrt{\left(1 - \frac{c_{13}^2}{c_{11} \times c_{33}}\right)}, \quad (3)$$

where x_1 is the direction normal to the plate,[32,33] and v_{axial} the velocity of the compressional bulk wave propagating along axis x_3. Equation (3) applies to both isotropic and anisotropic bone material, and leads to the following values, according to Table I:

$$v_{S_0(\text{isotropic})} = 3215 \ \text{m s}^{-1},$$

$$v_{S_0(\text{anisotropic})} = 3560 \ \text{m s}^{-1}.$$

In both isotropic[18,19] and anisotropic cases, the lower limit of the observed range of SOS values therefore corresponds to the theoretical velocity of the nondispersive low frequency S_0 Lamb mode. Taking into account the acoustical anisotropy of cortical bone resulted in a significant reduction of the range of variation of SOS values as a function of cortical thickness from approximately 800 m s^{-1} for isotropic bone to approximately 450 m s^{-1} for transversely isotropic bone. It is worthy to note that *in vivo* studies[5,17,31] report variations in SOS values often larger than 700 m s^{-1} for several skeletal sites, which suggest that material properties and microporosity are additional important determinants of the observed variance of SOS values. The possibility of independent geometry and material properties measurements will be discussed later in Sec. VII.

VI. SENSITIVITY OF AXIAL TRANSMISSION TO CORTICAL POROSITY

Geometry and anisotropy effects were discussed in the two previous sections, considering a homogeneous material with fixed material properties. However, both material properties at the tissue level and porosity are determinants of SOS. During aging and osteoporosis, cortical resorption results in an increased porosity in the inner part of the cortical shell.[10–12] The question is to know to what extent ultrasonic waves propagating at the surface of the bone may reflect changes in porosity occurring mainly in the endosteal region. In this section, our goal is to assess the sensitivity of the technique to changes in cortical porosity, for fixed material properties. We first measure effective velocity values as a function of porosity in a homogeneously porous medium, using plane waves. Using gradients of effective material properties, we then assess the sensitivity of axial transmission measurements to gradient of porosity increasing from the periosteal to the endosteal region.

A. Effective velocities as a function of cortical porosity

1. Objectives

Here, only pores corresponding to Haversian canals observed in Fig. 5 are taken into account. As this intracortical porosity involves pores dimensions of the order of 50–200 μm,[26] a comprehensive and accurate 3D simulation of ultrasonic propagation in such a porous cortical shell requires a simulation spatial pitch of the order of a few tens of microns. 3D simulation domains of a few cm^3, needed to simulate a complete axial transmission velocity measurement, would thus lead to huge memory and computational time requirements, making such simulations unrealistic today. Given the small size of the pores compared to the wavelength, porous bone acts locally as an effective homogeneous elastic medium. So, the idea was first to model porous cortical bone as an effective medium, whose properties can then be used as initial guess in the model used to simulate the axial transmission SOS measurement. To derive effective ultrasonic velocity values, we performed 3D simulations of ultrasonic transmission of plane waves through homogeneously porous models as described in Sec. III B. The spatial pitch was set to a value of 20 μm. Preliminary comparisons of results obtained on periodic and random porosity networks showed no significant difference between effective velocity values measured for these different types of networks. Periodic networks were then preferred in the complete set of simulation, since symmetry and invariance properties of such periodic porosity allowed fast 2D and 3D simulations, strictly equivalent to simulations performed in unbounded media transversely to the direction of propagation. Planes waves were propagated both along and across the axis of the cylindrical pores, in order to derive transverse and axial effective compressional bulk wave velocities. The simulation dimension along the direction of propagation was approximately three compressional wavelengths, with PML conditions on the two boundaries encountered by the plane wave. 500 kHz, 1 MHZ, and 2 MHz broadband pulses were transmitted and time delays needed to propagate through the whole simulation domain were used to derive effective velocities. Such simulated measurements are analogous to experimental measurements performed by Yoon and Katz,[28] which led to the values reported in Table I.

2. Results

Results are plotted in Fig. 12. Compressional velocities in both directions are found to be independent of frequency. Compressional bulk wave velocity appears to be approximately twice more sensitive to porosity when the wave propagates perpendicularly to pores axis (in the radial direction of bone) compared to the propagation parallel to pores axis (along the axis of bone). Typically, a 15% porosity leads

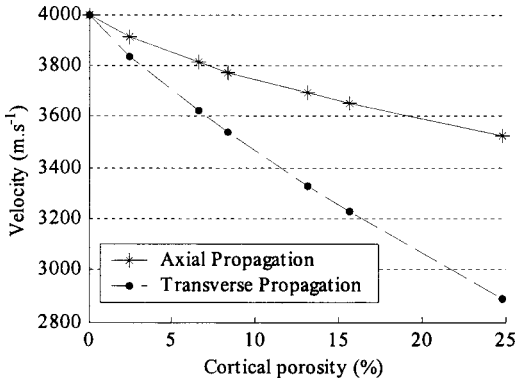


FIG. 12. Effective compressional bulk wave velocity as a function of cortical porosity, for propagation along (axial) and across (transverse) the pores axis.

to a 300 $m\,s^{-1}$ decrease of the compressional velocity along pores axis and a 600 $m\,s^{-1}$ decrease across the pores axis, relative to the homogeneous case. This shows that the acoustical anisotropy of bone is enhanced by the anisotropy of the porosity, in addition to the anisotropy due to structural organization at the osteons level.[28]

B. Cortical depth contributing to axial transmission measurement

In this part, our goal is to assess the cortical depth that influences the SOS value measured in axial transmission. Indeed, bone mechanical properties are heterogeneous, in particular as porosity increases from periosteal to endosteal regions.[26] As previously discussed in Sec. VI A 1, 3D simulations of axial transmission measurements can only be carried out on effective homogenized bone. Following results from Sec. VI A, a positive porosity gradient from periosteal to endosteal regions can be considered equivalent to a negative effective velocity gradient. Although effective velocities are different in axial and transverse directions as previously discussed, we used locally isotropic constants as a first approach to transform the porosity gradient into an isotropic velocity gradient. A typical porosity gradient value of 5%/mm[26] was modeled by a $-100\ m\,s^{-1}$/mm compressional effective velocity gradient, with a 4000 $m\,s^{-1}$ velocity at the periosteal interface. To avoid confounding thickness effects,[18] the gradient extended over a depth larger than three wavelengths, for each simulated frequency. This part of our study is therefore limited to the sensitivity of the *lateral* wave velocity to cortical porosity. The lateral wave velocity was measured at different locations along the interface, for different frequencies (500 kHz, 1 MHz, and 2 MHz). To quantitatively assess the sensitivity to the porosity, we report absolute SOS values, as well as the difference (noted ΔSOS) between the reference SOS of 4000 $m\,s^{-1}$ and SOS values measured with the velocity gradient.

Curves plotted in Figs. 13(a) and (b) represent measured absolute SOS values and ΔSOS, as a function of emitter–receiver distance normalized to compressional wavelength (defined from the reference velocity of 4000 $m\,s^{-1}$). As a velocity gradient of $-100\ m\,s^{-1}$/mm was used, equivalent depths were associated to ΔSOS values on a second vertical

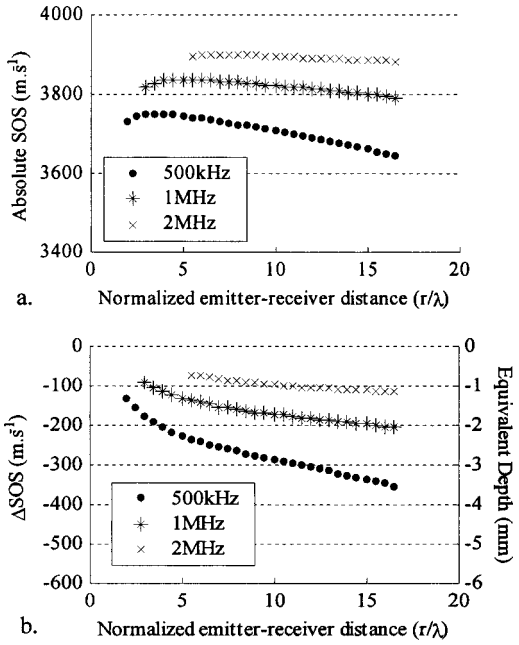


FIG. 13. Effect of a velocity gradient ($-100\ m\,s^{-1}$/mm) across cortical thickness on SOS as a function of emitter–receiver distance. Results are expressed as absolute values (a) and differences in SOS (b), relative to the case of homogeneous bone measurements.

axis [Fig. 13(b)], as a convenient and intuitive way to quantitatively assess the cortical depth probed by axial transmission SOS measurements. Figure 13(b) clearly shows that the sensitivity to porosity, expressed as ΔSOS values or equivalent depth, depends on frequency. Nevertheless, the emitter–receiver distance also has significant effects on the shape of curves plotted in Fig. 13(a). The increase in SOS values observed for small distances is due to near-field effects reported in Sec. II C, that cancel out for ΔSOS values [Fig. 13(b)]. For larger distances, increased ΔSOS values (i.e., increased equivalent depths) are observed, in agreement with distance effects reported in the case of homogeneous bone plates models:[18] the larger the emitter–receiver distance, the deeper the region that contributes to the first arriving signal. Thus, the sensitivity of axial transmission SOS to deep cortical porosity depends on both frequency and propagation distance. For instance, considering a ten wavelengths emitter–receiver distance (i.e., 8 cm at 500 kHz, 4 cm at 1 MHz, and 2 cm at 2 MHz), equivalent depths at 500 kHz, 1 MHz, and 2 MHz are, respectively, approximately 3, 1.75, and 1 mm. For this transducer's configuration, the magnitude of SOS variations due to deep cortical porosity significantly increases with increasing wavelength. The equivalent depth is approximately half a wavelength. The comparison of the sensitivity at several frequencies, for a fixed emitter–receiver distance, yields totally different conclusions: for a 20 mm emitter–receiver distance, which is a typical distance found in probes used *in vivo* to measure the radius,[1,6] equivalent depths are now 1.6 mm (500 kHz), 1.4 mm (1 MHz), and 1.1 mm (2 MHz). In this particular case, our results suggest that the porosity at depth around 3–4 mm probably has a minor or insignificant effect on the measured SOS.

It should be kept in mind that these results have been obtained on thick media, without any thickness effects in-

volved in the measurements. Curves in Figs. 13(a) and (b) can therefore only be used for emitter–receiver configurations that lead to measurements independent of cortical thickness. Such configurations have been detailed in a previous 2D study.[18] Therefore, the conclusions drawn here are valid only for the lateral wave.

C. Discussion

In Sec. VI A, porous bone is modeled by marrow inclusions in an isotropic bony matrix with fixed material properties, for different porosity values. Therefore, this approach does not take anisotropy into account at the materiel level. Our results show that the anisotropic porous network results in anisotropic effective elastic constants, even when one considers an isotropic bony matrix. For a 5% porosity (respectively, 10%), the corresponding relative difference $(V_{\text{axial}} - V_{\text{transverse}})/V_{\text{axial}}$ is approximately 4% (respectively, 8%). Experimentally, a relative difference of 15% between axial and transverse velocity values has been reported (cf. Table I). Our results therefore suggest that the elastic anisotropy at the material level contributes to the total effective elastic anisotropy of bone. Further work is needed to determine whether taking into account anisotropy of the material itself has a significant effect on the sensitivity to porosity. Moreover, intracortical porosity was modeled as a first step by cylindrical marrow inclusions, to account for porosity due to Haversian canals. A more refined model must be developed in order to take into account additional features of the actual porosity network, in particular endosteal resorption cavities.

In Sec. VI B, we used effective velocity values derived from Sec. VI A to simulate the effect of porosity gradients. This approach has several limitations. Porosity does not only affect velocity, but also has an impact on wave attenuation, at least through scattering mechanisms. Attenuation effects have therefore been neglected as porosity was only modeled by the corresponding changes in effective velocity values. In particular, interactions occurring at the boundary of the pores and the fluid–solid coupling are not taken into account. Moreover, the anisotropy of the effective elastic tensor was neglected. Nevertheless, as the relative difference $(V_{\text{axial}} - V_{\text{transverse}})/V_{\text{axial}}$ is of the order of 15% only, we expect results from Sec. VI B to be close to results that one would obtain with anisotropic effective constants. We think that our simplified approach yields a first reliable quantitative answer to the question of axial transmission sensitivity to cortical bone porosity, when a lateral wave is measured. A good sensitivity to endosteal porosity requires one to use a wavelength of the order of or larger than cortical thickness. In this case, the first detected signal does not correspond anymore to a lateral wave. Combined effects of cortical thickness and endosteal porosity may occur in this case, which were not assessed in this study.

VII. CONCLUSION

Ultrasound axial transmission through a skeletal site is complex and depends on a variety of skeletal parameters as well as on the experimental configuration. The research field has produced several experimental studies looking at the relationships of bone properties such as BMD and cortical thickness, but without elucidating the role played by elastic anisotropy, porosity, or macroscopic shape in wave propagation. An accurate interpretation of ultrasound measurements first requires a detailed understanding of ultrasound propagation with clear identification of the different waves that contribute to analyzed signals and their exact propagation path. It is virtually impossible to model by analytic means the extremely complex field resulting from the interaction of an incident wave with cortical bone taking into account the full complexity of the geometry, boundary conditions, anisotropy, and heterogeneity in bone properties. Recently developed simulation methods based on finite difference[24,34,35] offer a fertile alternative to inextricable analytic formulations. Such wave propagation simulation has been applied to the problem of axial transmission along the radius in the 2D case.[18,19,21] Numerical simulations have been found of great value in giving insight into the properties (nature, pathway) of propagating waves. It can be used to elucidate the relationship between SOS and bone properties (elasticity, porosity, cortical thickness) and test inverse calculation procedures. To the best of our knowledge, this study is the first study to investigate effects of 3D cortical bone geometry, anisotropy and heterogeneous porosity on axial transmission measurements.

We first showed that 2D plate models were equivalent to tubular models, in predicting axial transmission measurements based on the first arriving signal. This result validates the conclusions of previous 2D approaches.[18,19] We then showed that the anisotropy of cortical bone had a major impact on SOS values. The range of SOS values that may be expected from axial transmission measurements on different cortical thicknesses, for given anisotropic material properties, is approximately twice smaller than that one would obtain on isotropic bone models, and the dependence on thickness occurs in a narrower range of thicknesses, between $\lambda_{\text{bone(axial)}}/4$ and $\lambda_{\text{bone(axial)}}/2$. SOS not only depends on thickness, but also on material properties, i.e., elasticity and density. Therefore, an important issue is the feasibility of independent measurement of material and geometrical properties of bone. For small thickness to wavelength ratio (i.e., low frequency), the S_0 lamb mode velocity arrives first, whereas the axial compressional bulk wave velocity is measured for large thickness to wavelength ratio (i.e., high frequency). Both the compressional bulk velocity and S_0 mode velocity are independent of thickness, which suggests that an appropriate multiple-frequency approach could potentially yield measurements of material properties, independent of geometrical properties. Sensitivity to thickness may then be achieved for intermediate frequencies. For such intermediate frequencies however, SOS value measurements may be difficult to perform because signal amplitude has been reported to be relatively weak, with signal shapes and velocities highly dependent on emitter–receiver distances and signal processing.[18] Additional amplitude measurement may also provide valuable information, although we focused our interest on SOS measurement. Further work is needed to investigate the relevance of amplitude measurements. Axial and transverse effective velocity values as a function of intracor-

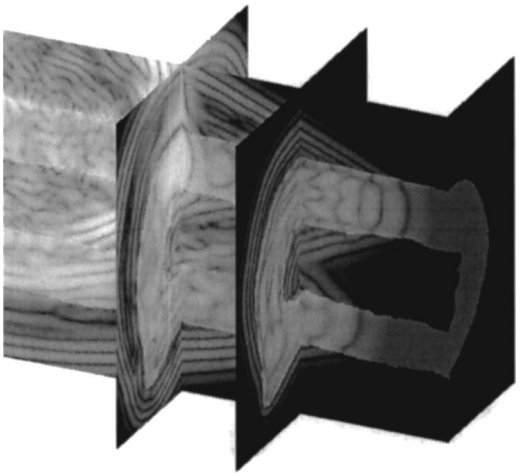


FIG. 14. 3D snapshot of ultrasonic waves propagating through a real cortical bone geometry obtained from 3D x-ray computerized tomography.

tical porosity have been assessed and used to model gradients of porosity. We showed that sensitivity to endosteal porosity may only be achieved for appropriate frequency and emitter–receiver distance.

Although our simulation was carried out here on idealized 3D objects (tubes), 3D numerical simulations can be performed on actual bone geometry, as measured from 3D x-ray computerized tomography (CT) for instance, combined with local effective elastic constants. Such simulations are currently being performed on 3D numerical models of bone, obtained from human radius previously assessed experimentally by axial transmission in our laboratory. As an illustration, a typical 3D snapshot obtained by entering into the software a real 3D CT reconstruction from a human radius is shown in Fig. 14. The possibility of applying the 3D finite-difference approach to actual bone structures, such as shown in Fig. 14, provides a valuable tool to the study of axial transmission along cortical bone. Comparison between numerical predictions and experimental measurements should finally allow a comprehensive understanding of axial transmission phenomena involved *in vivo*.

ACKNOWLEDGMENTS

The authors wish to thank Dr V. Bousson (Hôpital Lariboisière) for providing x-ray 3D data, Guillaume Libaude (Laboratoire d'Imagerie Paramétrique) for 3D sketches, and Ingrid Leguerney (Laboratoire d'Imagerie Paramétrique) for the acoustic microscopy image.

[1] R. Barkmann, E. Kantorovich, C. Singal, D. Hans, H. K. Genant, M. Heller, and C. C. Glüer, "A new method for quantitative ultrasound measurements at multiple skeletal sites," J. Clinical Densitometry **3**, 1–7 (2000).

[2] G. Lowet and G. Van der Perre, "Ultrasound velocity measurement in long bones: Measurement method and simulation of ultrasound wave propagation," J. Biomech. **29**, 1255–1262 (1996).

[3] A. J. Foldes, A. Rimon, D. D. Keinan, and M. M. Popovtzer, "Quantitative ultrasound of the tibia: A novel approach for assessment of bone status," Bone (N.Y.) **17**, 363–367 (1995).

[4] M. Weiss, A. Ben-Shlomo, P. Hagag, and S. Ish-Shalom, "Discrimination of proximal hip fracture by quantitative ultrasound measurement at the radius," Osteoporosis Int. **11**, 411–416 (2000).

[5] C. Njeh, I. Saeed, M. Grigorian, D. L. Kendler, B. Fan, J. Shepherd, M. McClung, W. M. Drake, and H. K. Genant, "Assessment of bone status using speed of sound at multiple anatomical sites," Ultrasound Med. Biol. **27**, 1337–1345 (2001).

[6] D. Hans, S. Srivastav, C. Singal, R. Barkmann, C. Njeh, E. Kantorovich, C. C. Glüer, and H. K. Genant, "Does combining the results from multiple bone sites measured by a new quantitative ultrasound device improve discrimination of hip fracture?," J. Bone Miner. Res. **14**, 644–651 (1999).

[7] C. Njch, N. Shaw, J. N. Gardner-Medwin, C. M. Boivin, and T. R. Southwood, "Use of quantitative ultrasound to assess bone status in children with juvenile idiopathic arthritis," J. Clinical Densitometry **3**, 251–260 (2000).

[8] M. H. Lequin, I. M. van der Sluis, M. M. van den Heuvel-Eibrink, W. J. Hop, R. R. van Rijn, S. F. de Muinck Keizer-Schrama, and C. van Kuijk, "A longitudinal study using tibial ultrasonometry as a bone assessment technique in children with acute lymphoblastic leukaemia," Pediatr. Radiol. **33**, 162–167 (2003).

[9] Y. Littner, D. Mandel, F. B. Mimouni, and S. Dollberg, "Bone ultrasound velocity curves of newly born term and preterm infants," J. Pediatr. Endocrinol. Metab. **16**, 43–47 (2003).

[10] N. M. Keshawarz and R. R. Recker, "Expansion of the medullary cavity at the expense of cortex in postmenopausal osteoporosis," Metab. Bone Dis. Relat. Res. **5**, 223–228 (1984).

[11] C. Njeh, D. Hans, T. Fuerst, C. C. Glüer, and H. K. Genant, *Quantitative Ultrasound. Assessment of Osteoporosis and Bone Status* (Dunitz, London, 1999).

[12] H. Ritzel, M. Amling, M. Posl, M. Hahn, and G. Delling, "The thickness of human vertebral cortical bone and its changes in aging and osteoporosis: A histomorphometric analysis of the complete spinal column from thirty-seven autopsy specimens," J. Bone Miner. Res. **12**, 89–95 (1997).

[13] S. C. Lee, B. S. Coan, and M. L. Bouxsein, "Tibial ultrasound velocity measured in situ predicts the material properties of tibial cortical bone," Bone (N.Y.) **21**, 119–125 (1997).

[14] M. R. Stegman, R. P. Heaney, D. Travers-Gustafson, and J. Leist, "Cortical ultrasound velocity as an indicator of bone status," Osteoporosis Int. **5**, 349–353 (1995).

[15] S. Prevrhal, T. Fuerst, B. Fan, C. Njeh, D. Hans, M. Uffmann, S. Srivastav, and H. K. Genant, "Quantitative ultrasound of the tibia depends on both cortical density and thickness," Osteoporosis Int. **12**, 28–34 (2001).

[16] C. Njeh, D. Hans, C. Wu, E. Kantorovich, M. Sister, T. Fuerst, and H. K. Genant, "An *in vitro* investigation of the dependence on sample thickness of the speed of sound along the specimen," J. Med. Eng. Phys. **21**, 651–659 (1999).

[17] H. Sievänen, S. Cheng, S. Ollikainen, and K. Uusi-Rasi, "Ultrasound velocity and cortical bone characteristics in vivo," Osteoporosis Int. **12**, 399–405 (2001).

[18] E. Bossy, M. Talmant, and P. Laugier, "Effect of bone cortical thickness on velocity measurements using ultrasonic axial transmission: A 2D simulation study," J. Acoust. Soc. Am. **112**, 297–307 (2002).

[19] P. Nicholson, P. Moilanen, T. Kärkkäinen, J. Timonen, and S. Cheng, "Guided ultrasonic waves in long bones: Modelling, experiment and application," Physiol. Meas **23**, 755–768 (2002).

[20] E. Bossy, M. Talmant, and P. Laugier, "Bi-directional axial transmission improves accuracy and precision of ultrasonic velocity measurement in cortical bone," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **51**(1), 71–79 (2004).

[21] E. Camus, M. Talmant, G. Berger, and P. Laugier, "Analysis of the axial transmission technique for the assessment of skeletal status," J. Acoust. Soc. Am. **108**, 3058–3065 (2000).

[22] L. Brekhovskikh and O. Godin, *Acoustics of Layered Media. II. Point Sources and Bounded Beams. Wave Phenomena* (Springer, Berlin, 1992).

[23] J. Virieux, "P-SV wave propagation in heterogenous media: Velocity-stress finite-difference method," Geophysics **51**, 889–901 (1986).

[24] R. W. Graves, "Simulating seismic wave propagation in 3D elastic media using staggered-grid finite differences," Bull. Seismol. Soc. Am. **86**, 1091–1106 (1996).

[25] F. Collino and C. Tsogka, "Application of the PML absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media," Geophysics **66**, 294–307 (2001).

[26] V. Bousson, A. Meunier, C. Bergot, E. Vicaut, M. A. Rocha, M. H. Morais, A.-M. Laval-Jeantet, and J.-D. Laredo, "Distribution of intracortical porosity in human midfemoral cortex by age and gender," Ann. Physiol. Anthropol. **16**, 1308–1317 (2001).

[27] R. B. Martin, D. B. Burr, and N. A. Sharkey, *Skeletal Tissue Mechanics* (Springer, New York, 1998).

[28] H. S. Yoon and J. L. Katz, ''Ultrasonic wave propagation in human cortical bone. II. Measurements of elastic properties and microhardness,'' J. Biomech. **9**, 459–464 (1976).
[29] S. S. Mehta and P. P. Antich, ''Measurement of shear-wave velocity by ultrasound critical-angle reflectometry (UCR),'' Ultrasound Med. Biol. **23**, 1123–1126 (1997).
[30] J. L. Katz, H. S. Yoon, S. Lipson, and R. Maharidge, ''The effects of remodeling on the elastic properties of bone,'' Calcif. Tissue Int. **36**, S31–S36 (1984).
[31] M. Weiss, A. Ben-Shlomo, P. Hagag, and M. Rapoport, ''Reference database for bone speed of sound measurement by a novel quantitative multisite ultrasound device,'' Osteoporosis Int. **11**, 688–696 (2000).
[32] I. M. Daniel, T. Liber, and R. H. LaBedz, ''Wave propagation in transversely impacted composite laminates,'' Exp. Mech. **19**, 9–16 (1978).
[33] R. B. Ashman, S. C. Cowin, W. C. Van Buskirk, and J. C. Rice, ''A continuous wave technique for the measurement of the elastic properties of cortical bone,'' J. Biomech. **17(5)**, 349–361 (1984).
[34] G. Luo, J. J. Kaufman, A. Chiabrera, B. Bianco, J. H. Kinney, D. Haupt, J. T. Ryaby, and R. S. Siffert, ''Computational methods for ultrasonic bone assessment,'' Ultrasound Med. Biol. **25**, 823–830 (1999).
[35] R. S. Schechter, H. H. Chaskelis, R. B. Mignona, and P. P. Delsanto, ''Real-time parallel computation and visualization of ultrasonic pulses in solids,'' Science **265**, 1188 (1994).

Ultrasonic relaxation due to inclusion complex of amino acid by β-cyclodextrin in aqueous solution

Takanori Fukahori, Sadakatsu Nishikawa,[a] and Kyohei Yamaguchi
Department of Chemistry and Applied Chemistry, Faculty of Science and Engineering, Saga University, Saga, 840-8502, Japan

(Received 8 January 2004; revised 12 February 2004; accepted 16 February 2004)

Ultrasonic absorption coefficients in the frequency range of 0.8–95 MHz were measured in aqueous solutions of L-methionine or L-norleucine (guest) in the presence and absence of β-cyclodextrin (β-CD, host) at 25 °C. A single relaxational absorption was observed only in the solutions containing the guest and host. The ultrasonic relaxation was attributed to a perturbation of a chemical equilibrium associated with an interaction between β-CD and the amino acid to form the host–guest complex. The kinetic and thermodynamic parameters in the system of L-norleucine with β-CD were determined from the concentration dependence of the relaxation frequency and the maximum absorption per wavelength. Because of the concentration independence of the relaxation frequency in L-methionine system with β-CD, the equilibrium constant and the standard volume change of the complexation reaction were estimated first from the concentration dependence of maximum absorption per wavelength, and subsequently the rate constants were calculated with the help of the estimated equilibrium constant and the observed relaxation frequency. The results obtained in this study were compared with those for systems of β-CD with other amino acids or alcohols having comparable hydrophobicity.
[DOI: 10.1121/1.1695432]

PACS numbers: 43.80.Jz, 43.80.Cs [FD] Pages: 2325–2330

I. INTRODUCTION

It is well established that velocity and absorption of ultrasonic waves are closely related to a change of molecular structure and a nature of an intermolecular interaction in liquid.[1] Cyclodextrins (CDs) are cyclic oligosaccharides consisting of glucopyranose units linked by α-1,4-glucosidic bonds. The shape is likened to a hollow truncated bucket of which the interior cavity is hydrophobic while the annular regions of both cavities are hydrophilic in character due to hydroxyl groups. The unique structure of CDs (hosts) can accommodate various molecules (guests) into the cavity through noncovalent bonding to form so-called host-guest complexes. The complexation reactions of CDs are extensively studied stationary and kinetically by various kinds of static and dynamic methods, which include spectrophotometry, calorimetry, NMR, density, stopped-flow, temperature-jump, and ultrasonic relaxation methods.[2–10] However, the inclusion phenomenon accompanying volume change does not occur in the case of all guest molecules; a certain degree of hydrophobicity is needed to enter into the cavity.[6,11] It is found interesting that guest molecules are appreciably recognized by host CDs due to differences of hydrophobicity, isomeric structures, functional groups, existence of charge, and so forth.[12–15] Thus, CDs are expected to be good chemical carrier materials. For this reason, the supramolecular chemistry of CDs has been extensively performed in many fields such as pharmaceutical and industrial applications, and, additionally, CDs can be useful enzyme models for their recognition abilities in biological chemistry.[2–5,7]

However, kinetic measurements have been less prevalent because many of the inclusion processes in the supramolecular systems occur on very fast time scale. Ultrasonic relaxation method is one of the best tools to unveil the dynamic properties in liquids and solutions. Therefore, we have been especially focusing on the kinetics of recognition by CDs in aqueous media. The aim of the present study is to reveal how much a subtle structural difference in the hydrophobic part of amino acids can be recognized by β-CD from the kinetic viewpoint. For this purpose, L-methionine and L-norleucine that have a hydrophobic normal chain were chosen as guests for β-CD. The results obtained are compared with other amino acids that have branched carbon chain (L-leucine and L-isoleucine) and with several alcohols.

II. EXPERIMENTS

A. Chemicals

β-CD was purchased from Wako Pure Chemical Co. Ltd. and purified by recrystallization using distilled and filtered water by a Milli-Q SP-TOC filter system from Japan Millipore Ltd. L-methionine was also obtained from Wako Pure Chemical Co. Ltd. as purest grade and L-norleucine was from Aldrich Chem. Co. of which the purity is confirmed more than 99%. They were used without further purification. All of the sample solutions were prepared by weighing just before the experimental measurements. The experimental *p*Hs were at their isoelectric points of the amino acids studied.

[a]Author to whom correspondence should be addressed. Electronic mail: nishikas@cc.saga-u.ac.jp

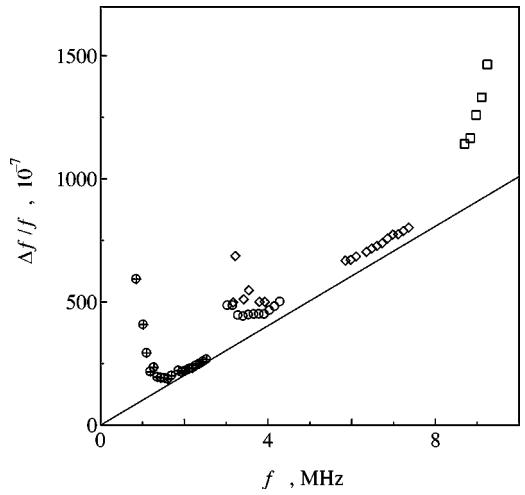


FIG. 1. Plots of $\Delta f/f$ versus ultrasonic frequency for water; (⊕): from 3 MHz cell, (○): 5 MHz cell A, (◇): 5 MHz cell B, and (□): 7 MHz cell. The solid line is drawn by using constant value of a/f^2 and sound velocity.

B. Apparatus

Ultrasonic absorption coefficients, α, were measured by a resonance method in the frequency range from about 0.8 to 9 MHz. The apparatus consists of four resonance cells with 3-, 5-, and 7-MHz fundamental x-cut crystals. The 3-MHz cell was used for the frequency range from about 0.8 to 2.5 MHz, the 5-MHz cell **A** covered from about 3 to 4.5 MHz, the 5-MHz cell **B** from 3 to 7.5 MHz, and the newly constructed 7-MHz cell was from 8- to 9-MHz range. The temperature control for the resonator cells was maintained within ± 0.01 °C (Lauda RM20). The coefficient, α, by the resonance method is obtained from the half-bandwidth of the resonance curves, Δf. Since the values of Δf are usually enlarged due to a mechanical loss, it was needed to measure the values of $\Delta f/f$ for water as a reference liquid over the measurement frequency prior to the experiments in order to check the extent of broadening. This was because the acoustic impedance of the solutions investigated was close to that for water. In water at 25 °C, the values $\alpha/f^2 = 21.1 \times 10^{-15}\,\mathrm{s^2\,m^{-1}}$ (f: ultrasonic frequency) and v (sound velocity) $= 1497\,\mathrm{m\,s^{-1}}$ were taken as constants over the measurement frequency range. Thus, the differences between the theoretical values and practical ones for water were obtained as the mechanical loss, Q^{-1}, by which α/f^2 values for solutions were calculated through the equation as $\alpha/f^2 = (\Delta f/f - Q^{-1})(\pi/v)/f$.[16] Figure 1 shows the experimental values for the individual cells and the calculated line for water using the above constant values of α/f^2 and v.

A pulse method with 5-MHz fundamental crystal was used in the range from 25 to 95 MHz and the temperature for the pulse cell was maintained within ± 0.1 °C (EYEYA UNI ACE BATH NCB-2200). Sound velocity values were obtained by the resonance method at around 3 MHz and solution densities were measured by using a vibrating density meter (Anton Paar NMA 60/602). The solution pH was measured by using a glass electrode with HM-60S Toa Denpa pH meter. All measurements were performed at 25 °C.

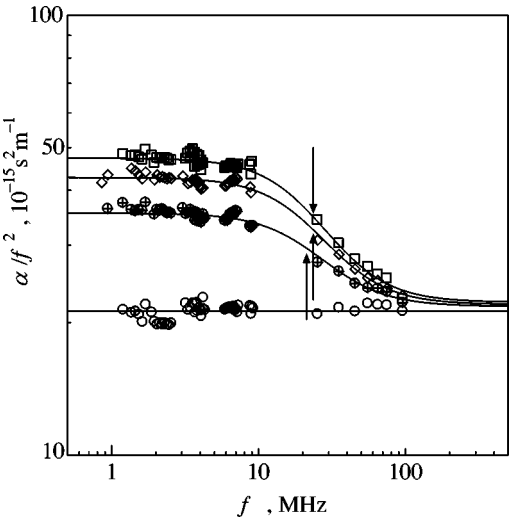


FIG. 2. Representative ultrasonic absorption spectra in aqueous solution of L-methionine in the presence and absence of 0.0087 mol dm^{-3} β-CD at 25 °C; (⊕): 0.040 mol dm^{-3} L-methionine+β-CD, (◇): 0.070 mol dm^{-3} L-methionine+β-CD, (□): 0.100 mol dm^{-3} L-methionine+β-CD, and (○): 0.100 mol dm^{-3} L-methionine only. The arrows show the locations of relaxation frequency.

III. RESULTS AND DISCUSSION

Figures 2 and 3 show representative ultrasonic absorption spectra in aqueous solution of L-methionine and L-norleucine in the presence and absence of β-CD. No additives such as acid, base, or buffer were used in the sample solutions; the amino acids were existing as zwitterions. The absorption coefficients divided by the square of the measurement frequency, α/f^2, are independent of the frequency when only the guest molecule (L-methionine or L-norleucine) was dissolved in water. Correspondingly, no excess absorption is observed in aqueous solution of β-CD below 0.01 mol dm^{-3}.[17] However, when the two solutes, β-CD and L-methionine or L-norleucine, coexisted in water

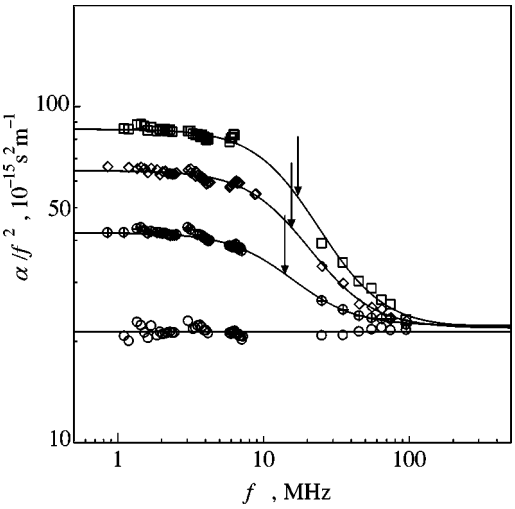


FIG. 3. Representative ultrasonic absorption spectra in aqueous solution of L-norleucine in the presence and absence of 0.0087 mol dm^{-3} β-CD at 25 °C; (⊕): 0.020 mol dm^{-3} L-norleucine+β-CD, (◇): 0.050 mol dm^{-3} L-norleucine+β-CD, (□): 0.090 mol dm^{-3} L-norleucine+β-CD, and (○): 0.090 mol dm^{-3} L-norleucine only.

TABLE I. Ultrasonic relaxation and thermodynamic parameters for aqueous solution of L-methionine and L-norleucine with β-CD at 25 °C.

C_{CD} (mol dm^{-3})	C_{GST} (mol dm^{-3})	f_r MHz	A (10^{-15} s^2 m^{-1})	B (10^{-15} s^2 m^{-1})	ρ kg m^{-3}	v m s^{-1}	pH
			L-methionine system				
0.0087	0.030	22.8±1.1	11.1±0.3	22.0±0.1	1002.23±0.01	1500.3±0.8	5.92
0.0087	0.040	20.7±0.6	13.5±0.2	22.3±0.1	1002.66±0.01	1500.9±0.8	5.92
0.0087	0.050	24.6±0.9	16.0±0.4	21.6±0.1	1003.07±0.01	1502.6±0.9	5.93
0.0087	0.070	23.8±0.4	20.8±0.2	22.0±0.1	1004.03±0.01	1504.0±0.7	6.02
0.0087	0.080	24.2±0.6	22.6±0.3	21.7±0.1	1004.45±0.01	1504.3±0.8	5.87
0.0087	0.090	24.6±0.5	24.7±0.3	21.8±0.1	1004.58±0.01	1505.8±0.8	5.89
0.0087	0.100	23.5±0.3	25.1±0.2	22.3±0.1	1005.15±0.01	1507.5±0.8	5.91
			L-norleucine system				
0.0087	0.020	13.4±0.3	19.8±0.3	22.3±0.1	1001.42±0.01	1498.4±0.7	6.02
0.0087	0.030	15.3±0.3	27.2±0.3	22.6±0.1	1001.61±0.01	1501.0±0.8	6.22
0.0087	0.040	16.0±0.4	34.2±0.4	22.2±0.1	1001.83±0.01	1501.8±0.8	6.14
0.0087	0.050	15.9±0.2	42.5±0.3	21.9±0.1	1002.08±0.01	1502.9±0.7	6.16
0.0087	0.080	16.7±0.3	59.9±0.6	22.2±0.1	1002.76±0.01	1506.1±0.9	6.17
0.0087	0.090	17.1±0.4	63.7±0.9	22.0±0.2	1003.01±0.01	1507.0±0.8	6.12

in the concentration range below which no excess absorption is observed individually, a relaxation phenomenon was clearly observed.

A Debye-type single relaxational equation was used to analyze the frequency dependence of the absorption as

$$\alpha/f^2 = A/\{1+(f/f_r)^2\}+B, \tag{1}$$

where f_r is the relaxation frequency, A is the amplitude of the excess absorption, and B is the background absorption. A nonlinear least-mean-squares method was applied to determine the above three ultrasonic parameters in order to obtain the best fit of the experimental values of α/f^2 to Eq. (1). The solid lines shown in Figs. 2 and 3 are drawn by using determined parameters, f_r, A, and B. It is obvious that the calculated line and experimental data agree very well. Table I shows the obtained ultrasonic relaxation parameters together with the solution density, ρ, the sound velocity, v, and the pH of each system.

It is clear that the cause of the relaxation is related to an interaction between β-CD and the guest molecule since the relaxation appears only when the solution contains the two solutes. Since only single relaxations were observed in the measurement frequency range in all solutions containing host and guest, the cause of the relaxation was considered to be responsible for a perturbation of the following chemical equilibrium by ultrasonic wave as follows,

$$\mathrm{CD}+\mathrm{GST}^{\pm} \underset{k_f}{\overset{k_b}{\leftrightarrows}} \mathrm{CDGST}^{\pm}, \tag{2}$$

where k_f is forward rate constant and k_b is the backward rate constant, and CD, $\mathrm{GST}^{\pm}$, and $\mathrm{CDGST}^{\pm}$ indicate β-CD as the host, amino acid as the guest, and host-guest complex, respectively. The relation between the relaxation time, τ, or the relaxation frequency, f_r, and the reactant concentrations can be derived as

$$\tau^{-1} = 2\pi f_r = k_f\{[\mathrm{CD}]+[\mathrm{GST}^{\pm}]\}+k_b \tag{3}$$

$$= k_b\{(KC_{CD}+KC_{GST}+1)^2-4K^2C_{CD}C_{GST}\}^{1/2}, \tag{3'}$$

where K is the equilibrium constant defined as $K=k_f/k_b$, and C_{CD} and C_{GST} are the initial concentration of the host and the guest, respectively. The initial concentration of guest, C_{GST}, is the only variable for the relaxation frequency when the concentration of β-CD is kept constant (i.e., $C_{CD}=0.0087$ mol dm^{-3} in this study). It is seen that the relaxation frequency increases with guest concentration (L-norleucine) as shown in Table I. Thus, the kinetic parameters, k_b and K, can be estimated from Eq. (3′) by using a nonlinear least-mean-squares method. Figure 4 shows the plots and calculated line of $2\pi f_r$ vs $\{(KC_{CD}+KC_{GST}+1)^2-4K^2C_{CD}C_{GST}\}^{1/2}$ using values of k_b and K obtained from Eq. (3′) for L-norleucine system. The good agreement between experimental data and the calculated line indicates that the estimated kinetic parameters are valid for the complexation reaction between β-CD and L-norleucine. They are listed in Table II along with those for other systems with β-CD for reference and discussion. However, this analytical procedure can be used only for systems in which the relax-

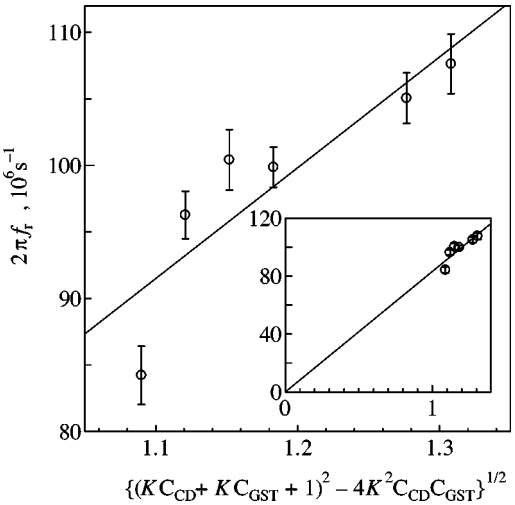


FIG. 4. Plots of $2\pi f_r$ vs. $\{(KC_{CD}+KC_{GST}+1)^2-4K^2C_{CD}C_{GST}\}^{1/2}$ for aqueous solution of L-norleucine in the presence of 0.0087 mol dm^{-3} β-CD at 25 °C.

TABLE II. The rate constants and standard volume change for host–guest complexation at 25 °C.

Guest	k_f (10^8 mol^{-1} dm^3 s^{-1})	k_b (10^7 s^{-1})	K (mol^{-1} dm^3)	ΔV (10^{-6} m^3 mol^{-1})
1-butanol	2.8±0.8	3.8±0.6	7.2±2.0	11.1±1.0
2-methyl-1-propanol	4.2±0.1	1.50±0.03	27.9±0.7	11.0±0.2
L-leucine	(2.6)	(7.9±0.3)	(3.3)	(16.3±2.7)
L-isoleucine	2.9±0.3	5.9±0.2	4.9±0.4	10±1
L-methionine	(3.7)	(14.8±0.6)	(2.5)	(15.4±3.5)
L-norleucine	2.7±0.2	8.3±0.1	3.3±0.2	18.6±0.4
butylammonium chloride	0.8±0.1	24.3±0.1	0.33±0.02	25.6±4.2

ation frequency is dependent on the guest concentration. Since the relaxation frequencies do not depend on the guest concentration in the system with L-methionine, another analytical procedure is needed.

Maximum absorption per wavelength, μ_{max}, is another important quantity that can be obtained from the ultrasonic absorption measurements as

$$\mu_{max}=0.5Af_r v=\pi\rho v^2\{1/[\mathrm{CD}]+1/[\mathrm{GST}^{\pm}]+1/[\mathrm{CDGST}^{\pm}]\}^{-1}(\Delta V)^2/2RT, \quad (4)$$

where ΔV is the standard volume change of the reaction, R is the gas constant, and T is the absolute temperature. The enthalpy term in Eq. (4) can be negligible because all the experiments were carried out in an aqueous solution. It is possible to calculate the individual equilibrium concentration of the reactants, once the equilibrium constant, K, is determined. Therefore, ΔV was calculated for L-norleucine system at the various concentrations of the guest, and the mean value is shown in Table II.

For the case of L-methionine system, the experimental results show that the guest concentration is almost independent of the relaxation frequency while the excess absorption, A, is increasing with the guest concentration. This result was also observed for the system of L-leucine and β-CD.[21] In order to obtain the equilibrium constant for L-methionine system, Eq. (4) was converted to the following form,

$$2RT\mu_{max}/\pi\rho v^2=\{1/[\mathrm{CD}]+1/[\mathrm{GST}^{\pm}]+1/[\mathrm{CDGST}^{\pm}]\}^{-1}(\Delta V)^2. \quad (5)$$

The left term in Eq. (5) can be obtained from experimental data. Therefore, various values of the equilibrium constant, K, were assumed for the plots of $2RT\mu_{max}/\pi\rho v^2$ vs $\{1/[\mathrm{CD}]+1/[\mathrm{GST}^{\pm}]+1/[\mathrm{CDGST}^{\pm}]\}^{-1}$ in order to obtain the best straight line going through a zero intercept. A value of $K=2.5\,\mathrm{mol}^{-1}\,\mathrm{dm}^3$ gave the best fit, as shown in Fig. 5. Then, the standard volume change was calculated from the slope of the line. As for the rate constants, Eq. (3) is approximated to $2\pi f_r \approx k_b$ on the assumption that the k_b value may be effectively greater than $k_f\{[\mathrm{CD}]+[\mathrm{GST}^{\pm}]\}$ term. Consequently, the k_b value was obtained from the average value of the observed relaxation frequencies and k_f from the definition of the equilibrium constant, K. The parameters obtained by this analytical procedure were used to confirm the above assumption. For example, the k_b value is approximately seven times greater than $k_f\{[\mathrm{CD}]+[\mathrm{GST}^{\pm}]\}$ term when $C_{GST}=0.050\,\mathrm{mol\,dm}^{-3}$. The parameters estimated through Eq. (5) are shown in parentheses in Table II without probable errors for k_f and K. In Fig. 5, similar plots for the L-norleucine system are shown in which the calculated K and ΔV values from Eqs. (3′) and (4) were used. The straight line going through the origin confirms that the relaxation is due to the complexation reaction between the host and the guest.

Hersey *et al.*[18,19] proposed that the complexation process of dyes to α-CD involves two steps. The first step is a fast preequilibrium one to form intermediate complex, followed by a slower rate-determining step to form final stable complex. They were detected by stopped-flow and temperature-jump methods and the kinetic parameters of second step were typically estimated on the millisecond time scale, which were very dependent on the nature of guests. Okubo *et al.*[20] examined an inclusion process between ionic detergent and CDs by a conductance stopped-flow method. The evaluated rate constants were also on the same time scale. However, when CDs complexation was studied by the ultrasonic relaxation method or a photoluminescence method,[8,21,22] the rate constants were from microsecond to nanosecond time scale and the forward rate constant was independent of the structure of host and guest while backward rate constant was very dependent. Despite only a single relaxation being observed in the present study, the possibility that another relaxation reflecting the second step complexation reaction might appear in considerably lower frequency

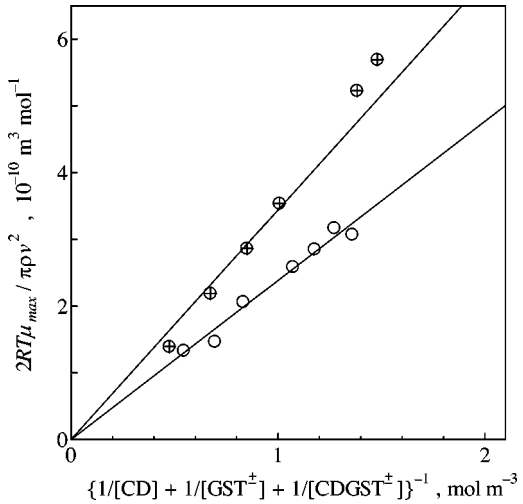


FIG. 5. Plots of $2RT\mu_{max}/\pi\rho v^2$ vs $\{1/[\mathrm{CD}]+1/[\mathrm{GST}^{\pm}]+1/[\mathrm{CDGST}^{\pm}]\}^{-1}$; (○): L-methionine system and (⊕): L-norleucine system.

range could not completely be excluded. However, the analysis for the present experimental results was proceeded following the one step reaction expressed by Eq. (2) as uncertain speculations complexify the kinetic analysis.

It was clarified in our previous examinations that the larger the hydrophobic part in the guest molecule is, the more firmly the guest molecule is incorporated into the cavity of β-CD. This is reflected in the smaller backward rate constant, k_b, and the greater equilibrium constant, K, while the forward rate constants, k_f, do not depend on the structure of the guest molecule, and, consequently, have similar values of about $3\times10^8\,\mathrm{mol}^{-1}\,\mathrm{dm}^3\,\mathrm{s}^{-1}$. The forward rate constant was calculated through Smoluchowski's equation, taking into account the necessary conditions for forming the inclusion complex, and the good agreement between calculated and experimental value was received.[21] Therefore, the inclusion process was confirmed to proceed through a diffusion-controlled reaction.

It is interesting to see how the backward rate constant, k_b, depends on the structure of the guest molecules. The amino acids series used as guests for β-CD is L-leucine, L-isoleucine, L-methionine, and L-norleucine, which have an analogous hydrophobic part. L-methionine is the molecule in which one of the methylene carbons in L-norleucine is replaced by a sulfur atom. Since unshared electron pairs exist on the sulfur atom in L-methionine molecule, it is predicted that the complex between L-methionine and β-CD might be less stable than that between L-norleucine and β-CD. Indeed, such a distinction in the molecular recognition is accomplished when the host is α-CD because the reported equilibrium constants for these two amino acids are considerably different; K is 8.9 $\mathrm{mol}^{-1}\,\mathrm{kg}$ for the L-methionine/α-CD system, and that for the L-norleucine/α-CD is 46 $\mathrm{mol}^{-1}\,\mathrm{kg}$.[23] However, the equilibrium constant obtained in the present study for the L-norleucine system is only slightly greater than for the L-methionine system, as is seen in Table II. The reason may be due to the difference of the cavity size, which is intimated by the experimental facts that the inclusion complex with α-CD is more stable than that with β-CD.[3,9] That is, the cavity size of β-CD may be too large to recognize precisely the difference between L-methionine and L-norleucine. On the other hand, it should be pointed out that the backward rate constant, k_b, in the L-methionine system is about twice as that for the L-norleucine system, which should reflect the difference in the hydrophobic portion of these two amino acids. L-leucine and L-isoleucine are the isomers of L-norleucine, having the branched hydrocarbon chain. It was also revealed that the guest molecules bearing branched chain are incorporated more snugly into the cavity when compared with those having normal chain. The fact leads to the smaller values of k_b and greater K and the trends can also be understood in alcohol systems.[12] From the results of present study, although they are not so intense, the values of k_b in the L-methionine and L-norleucine system are greater in comparison with the L-isoleucine or L-leucine system. It is concluded that the isomeric effect in amino acid on the inclusion complex reaction with β-CD can be principally seen in the backward rate constant, k_b, although they are not as prominent as is seen in C_3 or C_4 isomers of alcohol. In addition, it is also revealed that alcohol molecules can be included more firmly by β-CD than amino acid molecules bearing corresponding hydrophobic carbon chain. The fact that the pHs of the solutions correspond to the isoelectric points for these amino means that the amino acids are in their zwitterionic forms. The proximity of these charges to hydrophobic chain on the guest and to the hydroxyl groups at the annulus of the host of CD cavity would create a condition where penetration of the guest into the cavity would be difficult.

The following consideration arises from the results of the standard volume change of the reaction, ΔV. It was reported that the cavity of β-CD is able to accommodate approximately from five to seven water molecules although the exact number is not known.[24–26] When a guest molecule is incorporated into the cavity, a certain number of water molecules in high energy state are released to bulk phase.[3,6] When one methyl group and m methylene groups are incorporated into the host CD cavity and n water molecules are expelled from the cavity, the volume change of the reaction is simply derived as $\Delta V = nV_{\mathrm{H2O}} - mV_{\mathrm{CH2}} - V_{\mathrm{CH3}}$ in which $V_{\mathrm{CH2}} = 15.7\times10^{-6}\,\mathrm{m}^3\,\mathrm{mol}^{-1}$, and $V_{\mathrm{CH3}} = 27.1\times10^{-6}\,\mathrm{m}^3\,\mathrm{mol}^{-1}$.[6,14] It is only possible to receive the absolute value of the standard volume change from the ultrasonic absorption measurement. González-Gaitano *et al.*[26] precisely measured the solution density for surfactant and that with β-CD, from the result of which the volume change for complexation reaction was found to be positive.

It is well known that the hydrophobic group of the guest molecule can be incorporated into the β-CD cavity. If four methylene groups of norleucine were included into the cavity ($m=4$), then $n\approx6.0$; if the three of that were in that cavity ($m=3$), then $n\approx5.2$; and if $m=2$, then $n\approx4.3$, from the above relationship for the volume change. It is seen that these estimations are close to the butylammonium chloride and β-CD system.[14] It may be concluded that four or five water molecules on average are released from the β-CD cavity when the guest molecules studied in this report are incorporated into the β-CD cavity.

The obtained ΔV for the systems of amino acid with β-CD are relatively greater than those for the systems of compatible alcohols. One of the reasons may be due to the fact that amino acids could not enter into the hydrophobic cavity deeper compared to alcohol molecules, if the number of the expelled water molecules is similar. This is because of the zwitterionic form of amino acids. In fact, more increased volume change was obtained for butylammonium ion with β-CD system as is shown in Table II where the charge on guest molecule would cause an unfavorable condition for the penetration into the cavity.

[1] M. J. Blandamer, *Introduction to Chemical Ultrasonics* (Academic, London, New York, 1973).
[2] M. L. Bender and M. Komiyama, *Cyclodextrin Chemistry* (Springer-Verlag, New York, 1978).
[3] M. V. Rekharsky and Y. Inoue, "Complexation thermodynamics of cyclodextrins," Chem. Rev. (Washington, D.C.) **98**, 1875–1917 (1998).
[4] M. Fermeglia, M. Ferrone, A. Lodi, and S. Pricl, "Host-guest inclusion complexes between anticancer drugs and β-cyclodextrin: computational studies," Carbohydrate Polym. 1–30 (2003).
[5] S. Choudhury and A. K. Mitra, "Kinetics of aspirin hydrolysis and stabi-

lization in the presence of 2-hydroxypropyl-β-cyclodextrin,'' Pharm. Res. **10**, 156–159 (1993).
[6] L. D. Wilson and R. E. Verrall, ''A volumetric study of β-cyclodextrin/hydrocarbon and β-cyclodextrin/fluorocarbon surfactant inclusion complexes in aqueous solutions,'' J. Phys. Chem. B **101**, 9270–9279 (1997).
[7] R. Wimmer, F. L. Aachmann, K. L. Larsen, and S. B. Petersen, ''NMR diffusion as a novel tool for measuring the association constant between cyclodextrin and guest molecules,'' Carbohydr. Res. **337**, 841–849 (2002).
[8] N. J. Turro, T. Okubo, and C. J. Chung, ''Analysis of static and dynamic host-guest associations of detergents with cyclodextrins via photoluminescence methods,'' J. Am. Chem. Soc. **104**, 1794–1799 (1982).
[9] Y. Matsui and K. Mochida, ''Binding forces contributing to the association of cyclodextrin with alcohol in an aqueous solution,'' Bull. Chem. Soc. Jpn. **52**, 2808–2814 (1979).
[10] H. Fujiwara, H. Arakawa, S. Murata, and Y. Sasaki, ''Entropy changes in the inclusion complex formation of α-cyclodextrin with alcohols as studied by the calorimetry titration,'' Bull. Chem. Soc. Jpn. **60**, 3891–3894 (1987).
[11] T. Ugawa and S. Nishikawa, ''Kinetic study for molecular recognition of amino acid by cyclodextrin in aqueous solution,'' J. Phys. Chem. A **105**, 4248–4251 (2001).
[12] S. Nishikawa, T. Ugawa, and T. Fukahori, ''Molecular recognition kinetics of β-cyclodextrin for several alcohols by ultrasonic relaxation method,'' J. Phys. Chem. B **105**, 7594–7597 (2001).
[13] S. Nishikawa, T. Fukahori, and K. Ishikawa, ''Ultrasonic relaxations in aqueous solutions of propionic acid in the presence and absence of β-cyclodextrin,'' J. Phys. Chem. A **106**, 3029–3033 (2002).
[14] S. Nishikawa, K. Yamaguchi, and T. Fukahori, ''Ultrasonic relaxation due to complexation reaction between β-cyclodextrin and alkylammonium ions,'' J. Phys. Chem. A **107**, 6415–6418 (2003).
[15] Y. Liu, B. H. Han, A. D. Qi, and R. T. Chen, ''Molecular recognition study of a supramolecular system,'' Bioorg. Chem. **25**, 155–162 (1997).
[16] S. Kato, H. Nomura, R. Zieliński, and S. Ikeda, ''Precice measurement of ultrasonic absorption in micellar solutions in the frequency range from 0.2 to 10 MHz,'' Bull. Chem. Soc. Jpn. **59**, 707–713 (1986).
[17] S. Nishikawa and S. Yamaguchi, ''Kinetics of complexation between β-cyclodextrin and 1-propanol in aqueous solution by ultrasonic relaxation method,'' Bull. Chem. Soc. Jpn. **69**, 2465–2468 (1996).
[18] A. Hersey and B. H. Robinson, ''Thermodynamic and kinetic study of the binding of azo-dyes to α-cyclodextrin,'' J. Chem. Soc., Faraday Trans. 1 **80**, 2039–2052 (1984).
[19] A. Hersey, B. H. Robinson, and H. C. Kelly, ''Mechanism of inclusion-compound formation for binding of organic dyes, ions and surfactants to α-cyclodextrin studied by kinetic method based on competition experiments,'' J. Chem. Soc., Faraday Trans. 1 **82**, 1271–1287 (1986).
[20] T. Okubo, Y. Maeda, and H. Kitano, ''Inclusion process of ionic detergents with cyclodextrins as studied by the conductance stopped-flow method,'' J. Phys. Chem. **93**, 3721–3723 (1989).
[21] T. Fukahori, T. Ugawa, and S. Nishikawa, ''Molecular recognition kinetics of leucine and glycyl-leucine by β-cyclodextrin in aqueous solution in terms of ultrasonic relaxation,'' J. Phys. Chem. A **106**, 9442 (2002).
[22] D. Hall, D. Bloor, K. Tawarah, and E. Wyn-Jones, ''Kinetic and equilibrium studies associated with the formation of inclusion compound involving n-butanol and n-pentanol in aqueous cyclodextrin solutions,'' J. Chem. Soc., Faraday Trans. 1 **82**, 2111–2121 (1986).
[23] G. Barone, G. Castronuovo, V. D. Ruocco, V. Elia, and C. Giancola, ''Inclusion compounds in water: thermodynamics of the interaction of cyclomaltohexane with amino acids at 25 °C,'' Carbohydr. Res. **192**, 331–341 (1989).
[24] F. W. Lichtenthaler and S. Immel, ''On the hydrophobic characteristics of cyclodextrins: computer-aided visualization of molecular lipophilicity patterns,'' Liebigs Ann. 27–37 (1996).
[25] A. Marini, V. Berbenni, G. Bruni, and V. Massarotti, ''Dehydration of the cyclodextrins: a model system for the interactions of biomolecules with water,'' J. Chem. Phys. **103**, 7532–7540 (1995).
[26] G. González-Gaitano, A. Crespo, A. Compostizo, and G. Tardajos, ''Study at a molecular level of the transfer process of a cationic surfactant from water to β-cyclodextrin,'' J. Phys. Chem. B **101**, 4413–4421 (1997).

Potential sound production by a deep-sea fish

David A. Mann[a)]
University of South Florida, College of Marine Science. 140 7th Avenue S., St. Petersburg, Florida 33701

Susan M. Jarvis[b)]
Engineering Test and Evaluation Dept., Naval Undersea Warfare Center Division, Newport, Rhode Island 02841

(Received 2 November 2003; revised 6 February 2004; accepted 10 February 2004)

Swimbladder sonic muscles of deep-sea fishes were described over 35 years ago. Until now, no recordings of probable deep-sea fish sounds have been published. A sound likely produced by a deep-sea fish has been isolated and localized from an analysis of acoustic recordings made at the AUTEC test range in the Tongue of the Ocean, Bahamas, from four deep-sea hydrophones. This sound is typical of a fish sound in that it is pulsed and relatively low frequency (800–1000 Hz). Using time-of-arrival differences, the sound was localized to 548–696-m depth, where the bottom was 1620 m. The ability to localize this sound in real-time on the hydrophone range provides a great advantage for being able to identify the sound-producer using a remotely operated vehicle. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1694992]

PACS numbers: 43.80.Ka, 43.30.Sf [WA] Pages: 2331–2333

I. INTRODUCTION

Very little information is available on reproduction in deep-sea fishes. Most data has been obtained by making gonadosomatic index (i.e., gonad weight/total body weight) measurements of fishes from trawls (e.g., Pankhurst *et al.*, 1987; Bell *et al.*, 1992; D'Onghia *et al.*, 1999). These data do not provide much resolution on temporal or spatial scales of the reproductive patterns of deep-sea fishes. For instance, little is known about where deep-sea fishes spawn, whether they spawn in aggregations, and whether spawning is a continuous process or whether it occurs at only certain parts of the year.

Many deep-sea fishes are caught either intentionally (e.g., grenadier) or as bycatch and used for fish oil and other fish products (FAO, 2001). For example, 296 million kg of the Patagonian grenadier (*Macruronus magellanicus*) were harvested in 2001, making it the 40th most caught species by weight in the world. The next most common deep-sea species (43rd of all species harvested) was the blue grenadier (*Macruronus novaezelandiae*), of which 257 million kg was caught. Thus, deep-sea fishes represent important fisheries for which we have almost no information on their reproductive biology. It is important to understand reproductive dynamics in the deep-sea to be able to better predict the impact of fisheries on the deep-sea ecosystem.

Sound may be the best mode of communication in a large dark environment like the deep-sea because it is directional and can travel long distances. A number of benthopelagic deep-sea fishes, such as the Macrouridae (grenadiers or rattails), Ophidiidae (cusk eels), and Trachichthyidae (roughies) have sound-producing muscles on their swimbladders (Marshall, 1967; Shimizu, 1977). These muscles are most likely used to produce sound during courtship and spawning, as has been shown for a shallow-water cusk-eel (Mann *et al.*, 1997). We present the first description of a sound that is likely produced by a deep-sea fish, which is an important first step in being able to identify the actual species producing the sound.

II. METHODS

Recordings were made from four Atlantic Undersea Test and Evaluation Center (AUTEC) hydrophones, which are located off Andros Island, Bahamas. The hydrophones were located at 1620-m depth and spaced approximately 2 nm apart in a square pattern. Archived recordings were analyzed to cover a 24-h time period and an entire year (Table I). Each recording was 5.0 min long and sampled at 48 kHz. The potential fish sound was initially identified by listening. An automatic signal detection program was then written in MATLAB to isolate the calls from the recordings by comparing sound levels in the 800–1100-Hz band, where the sound had maximum energy, to sound levels in the 1200–1400-Hz band. This allowed both detection of signals of interest, and rejection of broad-band sperm whale click sounds, which were common in the recordings. The signal detection program performed the following steps: (1) band-pass filter with an 800–1100-Hz elliptic filter, (2) rectify, (3) low-pass filter with a 200-point FIR filter (moving average), and (4) compare this signal to a similarly processed signal that was band-pass filtered from 1200 to 1400 Hz. Signals that exceeded the threshold were then inspected to ensure that they were the same type of signal.

Time-of-arrival differences were calculated by performing pairwise cross correlations on the detected signals. These time-of-arrival differences were then used to calculate the location of the sound (see Jarvis and Moretti, 2002; Jarvis *et al.*, 2002). Not all signals had sufficiently high signal-to-noise ratio on each hydrophone to determine a three-dimensional location.

[a)]Electronic mail: dmann@marine.usf.edu
[b)]Electronic mail: jarvisSM@npt.nuwc.navy.mil

TABLE I. Summary of recordings analyzed for the presence of potential deep-sea fish sounds and whether a sound was detected. The time indicates the time of each 5.0-min recording.

Date	Time	Sounds detected
6-29-2000	00:30, 01:00	No
6-29-2000	02:30, 3:00	Yes
6-29-2000	05:30, 06:00	No
6-29-2000	08:30, 09:00	No
6-29-2000	11:30, 12:00	No
6-29-2000	14:30, 15:00	No
6-29-2000	17:00, 17:30	Yes
6-29-2000	20:00, 20:30	Yes
6-29-2000	23:00, 23:30	Yes
7-1-2000	00:00, 00:30	No
7-1-2000	12:00, 12:30	Yes
9-4-2000	23:29, 23:59	Yes
9-5-2000	11:00, 11:30	No
11-01-2000	00:00, 00:30	No
11-01-2000	12:00, 12:30	No
1-8-2001	00:00, 00:30	No
1-8-2001	12:00, 12:30	No

III. RESULTS

Sounds were present in recordings made in June, July, and September 2000 and were recorded during both day and night (Table I). A spectrogram of the recorded signal indicates a typical fishlike sound in that it is pulsed and low in frequency [Fig 1(a)]. Analysis of 23 sounds shows that they contain between 7 and 46 pulses (32 ± 9; mean$\pm$s.d.) with a pulse period of 25.0 ± 1.7 ms (mean$\pm$s.d.) [Fig 1(b)]. The mean peak frequency of these sounds was 901 ± 65 Hz. A strong surface echo is evident in all recordings, which often overlaps the direct path signal.

The sounds were localized in three dimensions using time-of-arrival information on four hydrophones spaced at approximately 4 km. Sounds with sufficient signal-to-noise ratio on all four hydrophones were recorded only on 1 July 2000 from 12:00 to 12:30. These were localized to 548–696-m depth ($n=5$), whereas the bottom depth was approximately 1600 m (Table II). Interestingly, each successive call was localized to a deeper depth. Assuming that the same individual made these calls, the descent rate would be 0.11 m/s.

The system on which these sounds were recorded has been dismantled, and precise calibration of this system is not known. Therefore, source levels can not be calculated directly. Source levels were estimated by comparing the signal-noise ratio (SNR) of received levels and comparing these to long-term measurements of ambient noise levels. Historical ambient levels for the Tongue of the Ocean between 800 and 1000 Hz are approximately 50 dB *re* 1 μPa/sqrt(Hz) for sea state 1. Sea state 1 was chosen based on historical weather data from this recording period, however the ambient noise level in the band of interest can range from 38.2 dB *re* 1 μPa/sqrt(Hz) for sea state 0 to 61.5 dB *re* 1 μPa/sqrt(Hz) for sea state 3. The SNR of the five localized sounds was 7.4 ± 2.9 dB (mean$\pm$SD), suggesting received levels of approximately 57 dB *re* 1 μPa/sqrt(Hz). The distance from the hydrophone to these five sounds is approximately 4404 ± 56 m (mean$\pm$SD). Assuming spherical spreading loss, the estimated source level for these five sounds (SL = Ambient Noise + SNR + 20 log(Distance)) is 131 ± 3 dB *re* 1 μPa/sqrt(Hz).

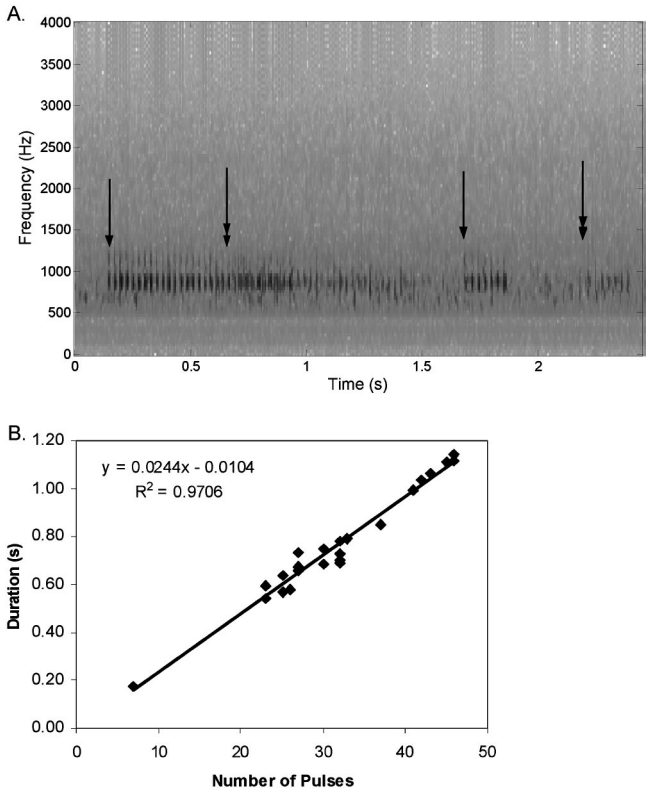


FIG. 1. (a) Spectrogram of putative deep-sea fish sounds recorded on a hydrophone located at approximately 1600-m depth in the Tongue of the Ocean, Bahamas. Two sounds are shown. A black arrow indicates the start of each sound. A double-arrow indicates the start of the echo (the echo overlaps with the direct path for the first sound). Sounds were originally digitized at 48 kHz, and downsampled to 8 kHz. A high-pass filter was applied to reduce noise below 500 Hz. (b) Plot of number of pulses in a sound versus sound duration with a linear regression of the data.

IV. DISCUSSION

There are no published data on sounds produced by deep-sea fishes. This paper is an attempt to propose a candidate deep-sea fish sound, which could then be used as a starting point for identifying the species that is producing this sound using a ROV or other method. Thus, it is necessarily speculative. The sounds presented in this paper are typical of sounds produced by fishes with sonic muscles on their swimbladders (e.g., Mann *et al.*, 1997). They are pulsed, stereotypical sounds, and the estimated source levels

TABLE II. Localization of five sounds on 1 July 2000 from the AUTEC hydrophone range. *X* and *Y* are presented in lieu of latitude and longitude coordinates (positive *Y* is North). Rate of movement is calculated assuming that the same individual made each call.

Time (s after 12:00)	*X* (m)	*Y* (m)	Depth (m)	Rate (m/s)
13.83	658	−630	−548	
164.03	651	−644	−562	0.14
263.45	656	−648	−569	0.10
1882.61	561	−733	−691	0.11
1998.44	549	−736	−697	0.12

are close to other fish sounds. There are several candidate species with sonic muscles attached to their swimbladders that may have produced these sounds. The most likely candidates are the deep-sea cods (Moridae), grenadiers (Macrouridae), cusk-eels (Ophidiidae), and roughies (Trachichthyidae) (FAO, 1990; Marshall, 1973; Shimizu, 1977). Since most fishes with sonic muscles produce sounds during courtship and spawning, positive identification of sound-producing species would allow the use of passive acoustics to study the life history of reproduction in these deep-sea fishes. For most of these fishes the seasonality and location of spawning is not known. While the data were insufficient to determine if there is seasonality in sound production, it is clear that the sound was made during both day and night.

The sounds produced by fishes during courtship and spawning can be used to determine the location of the spawning fish and to determine when the fish spawn. This technique has been successfully used with damselfishes to determine the periodicity of reproduction (Mann and Lobel, 1995) and sciaenids (croakers and drums) to identify areas used in spawning (Mok and Gilmore, 1983; Luczkovich *et al.*, 1999). Passive detection of sounds produced by deep-sea fishes could produce data on the timing and location of spawning that would be otherwise unobtainable. Many of these species are benthopelagic, and, while usually found near the bottom, can also be captured mid-water. The ability to localize this sound in real-time on the hydrophone range provides a great advantage for being able to identify the sound-producer using a remotely operated vehicle. Data on the positions that the sounds were localized to suggest that they might all come from the same individual, since the rate of ''movement'' was slow, relatively constant and in the same direction. Of course, it is impossible to determine whether these were calls from the same individual and/or multiple individuals. However, once individuals are identified, if there are individual differences in sound production, this could be a technique to study individual behavior.

One other possible source of this sound is from cetaceans, which are commonly recorded on the AUTEC range. However, most known cetacean sounds are either higher or lower in frequency than this sound. A pulsed sound between 800 and 1100 Hz would be expected for a deep-sea fish using sonic muscles on its swimbladder. Furthermore, fish sounds often have species-specific pulse rates that vary little between calls.

Beaked whales prey upon squids and deep-sea fishes (Walker *et al.*, 2003). Many of the deep-sea fishes that they prey upon, including the Moridae and Macrouridae, have sonic muscles on their swimbladders. Just as it has been hypothesized that bottlenose dolphins eavesdrop on sounds produced by prey fishes (Barros and Wells, 1998), it is reasonable to hypothesize that beaked whales and other cetaceans may use the same tactic in the deep-sea. The sounds recorded in the hydrophone range could be detected over a 4+km range, and thus may provide a reliable directional signal for cetaceans locating prey in the deep-sea. Furthermore, the delay from the surface echo could be useful for a deep-sea fish to determine its depth, and could also aid in localization by potential mates.

ACKNOWLEDGMENTS

SJ would like to acknowledge Dr. Robert Gisiner of ONR for funding the Marine Mammal Monitoring on Navy Ranges program, and also David Moretti, Dr. Ron Morrissey, Nancy DiMarzio and Jessica Ward of NUWC, Division Newport, and Eric Strahan of AUTEC, West Palm Beach for their contributions to the AUTEC data recording/localization efforts.

Barros, N. B., and Wells, R. S. J. (**1998**). ''Prey and feeding patterns of resident bottlenose dolphins (*Tursiops truncatus*) in Sarasota Bay, Florida,'' Mammol. **79**, 1045–1059.

Bell, J. D., Lyle, J. M., Bulman, C. M., Graham, K. J., Newton, G. M., and Smith, D. C. (**1992**). ''Spatial variation in reproduction and occurrence of non-reproductive adults, in orange roughy, *Hoplostethus atlanticus* Collett (Trachichthyidae), from south-eastern Australia,'' J. Fish Biol. **40**, 107–122.

D'Onghia, G., Basanisi, M., Matarrese, A., and Megli, F. (**1999**). ''Reproductive strategies in macrourid fish: seasonality or not?'' Mar. Ecol.: Prog. Ser. **184**, 189–196.

FAO Species Catalogue (**1990**). *Volume 10. Gadiform fishes of the world*, Rome, Italy.

FAO (**2001**). Fisheries Statistics, UN Food and Agriculture Organization.

Jarvis, S. M., and Moretti, D. J. (**2002**). Proc. Undersea Def. Tech. Europe, La Spazia, Italy, June 2002.

Jarvis, S., Moretti, D., Morrissey, R., and DiMarzio, N. (**2002**). ''Passive detection and localization of marine mammals in open ocean environments using widely spaced bottom mounted hydrophones,'' J. Acoust. Soc. Am. **114**, 2405.

Luczkovich, J. J., Sprague, M. W., Johnson, S. E., and Pullinger, R. C. (**1999**). ''Delimiting spawning areas of weakfish, *Cynoscion regalis* (Family Sciaenidae), in Pamlico Sound North Carolina using passive hydroacoustic surveys,'' Bioacoustics **10**, 143–160.

Mann, D. A., and Lobel, P. S. (**1995**). ''Passive acoustic detection of sounds produced by the damselfish, *Dascyllus albisella* (Pomacentridae),'' Bioacoustics **6**, 199–213.

Mann, D. A., Bowers-Altman, J., and Rountree, R. A. (**1997**). ''Sounds produced by the striped cusk-eel *Ophidion marginatum* (Ophidiidae) during courtship and spawning,'' Copeia 610–612.

Marshall, N. B. (**1967**). ''Sound-producing mechanisms and the biology of deep-sea fishes,'' in *Marine Bio-Acoustics. Volume 2*, edited by W. N. Tavolga (Pergamon, Oxford), pp. 123–133.

Marshall, N. B. (**1973**). ''Family Macrouridae,'' in *Fishes of the Western North Atlantic*, edited by D. M. Cohen, Mem. Sears Found. Mar. Res. no. 1 pt. 6. pp. 496–665.

Mok, H.-K., and Gilmore, R. G. (**1983**). ''Analysis of sound production in estuarine aggregations of *Pogonias cromis*, *Bairdiella chrysoura*, and *Cynoscion nebulosus*,'' Bull. Inst. Zool. Acad. Sinica **22**, 157–186.

Pankhurst, N. W. (**1988**). ''Spawning dynamics of orange roughy, *Hoplostethus atlanticus*, in mid-slope waters of New Zealand,'' Environ. Biol. Fish. **21**, 101–116.

Pankhurst, N. W., McMillan, P. J., and Tracey, D. M. (**1987**). ''Seasonal reproductive cycles in three commercially exploited fishes from the slope waters off New Zealand,'' J. Fish Biol. **30**, 193–211.

Shimizu, T. (**1977**). ''Comparative morphology of the expanded epipleural and its associated structures in four species of the Trachichthyidae,'' Jpn. J. Ichthyol. **23**, 192–198.

Walker, W. A., Mead, J. G., and Brownell, Jr., R. L. (**2003**). ''Diets of Baird's beaked whales, *Berardius bairdii*, in the southern Sea of Okhotsk and off the Pacific coast of Honshu, Japan,'' Marine Mammal Sci. **18**, 902–919.

A simple frequency-scaling rule for animal communication

Neville H. Fletcher[a)]
Research School of Physical Sciences and Engineering, Australian National University, Canberra 0200, Australia

(Received 6 November 2003; revised 27 January 2004; accepted 11 February 2004)

Different animals use widely different frequencies for sound communication, and it is reasonable to assume that evolution has adapted these frequencies to give greatest conspecific communication distance for a given vocal effort. Acoustic analysis shows that the optimal communication frequency is inversely proportional to about the 0.4 power of the animal's body mass. Comparison with observational data indicates that this prediction is well supported in practice. For animals of a given class, for example mammals, the maximum communication distance varies about as the 0.6 power of the animal's mass. There is, however, a wide spread of observed results because of the different emphasis placed upon vocal effort in the evolution of different animal species. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1694997]

PACS numbers: 43.80.Ka, 43.80.Lb [WA] Pages: 2334–2338

I. INTRODUCTION

Different animals employ widely different frequencies for their sound communication—communication that has the purpose of defining territory, seeking a mate, warning of danger, or simply social interaction. It is reasonable to suppose that evolution has led to the use of a frequency that gives greatest conspecific communication distance in at least the former cases, but what is the rule that governs this choice of frequency?

A detailed description of acoustic communication in a wide variety of animals has been given in two excellent books by Stebbins[1] and by Bradbury and Vehrenkamp,[2] and also in two classic collections of papers edited by Busnel[3] and by Lewis,[4] respectively. These books contain a wealth of information on many different animal species and give copious references to the original literature. Bradbury and Vehrenkamp also provide a clear, qualitative discussion of the physical processes underlying sound production, propagation, and hearing. A book by the present author,[5] in contrast, concentrates on the physical principles involved, and attempts to provide a quantitative mathematical basis upon which an understanding of individual cases can be built. A condensed account has been published elsewhere.[6]

Noting that small animals generally use higher frequencies in their calls than larger animals, one might conjecture a rule stating that call frequency is inversely proportional to the linear size L of the animal ($f \propto 1/L$ or $f \propto M^{-1/3}$, where M is the mass of the animal). Such a rule was proposed by Bradbury and Vehrenkamp[2] and justified by the observation that the radiation efficiency of an opening such as the mouth increases with increasing frequency until it reaches saturation when the wavelength of the sound is comparable with the diameter of the mouth (actually about equal to $\pi/2$ times the diameter[5]). While on the whole this suggests rather higher frequencies than are actually used, the fit to the observational data is quite good, as will be shown later. It is the purpose of the present note to derive a refinement of this rule on the basis of the physical principles underlying sound production, propagation, and hearing. The anatomic variety, lifestyle, and habitat of animals are, of course, immensely diverse so that a wide variation is to be expected, but perhaps the analysis will add to our understanding.

Most vertebrate land-dwelling animals produce sound by expelling air through a vibrating valve in the larynx, or syrinx in the case of birds. This valve leads to the upper respiratory tract, where air-column resonances, or ''formants,'' modify the spectral envelope of the richly harmonic sound generated by the oscillating valve, and sound is finally radiated through a mouth or beak, the diameter of which generally scales approximately as the linear size of the animal. Vocal information is generally encoded in the lowest two or three formant bands—regions of emphasis in the overtone spectrum—extending up to about 10 times the fundamental frequency. There is an exception to this statement in the case of many passerine birds and some frogs, which produce calls consisting essentially only of a fundamental with all its harmonics suppressed.

Insects, on the other hand, must produce sound by purely mechanical means, since they have no respiratory air supply under pressure. This is generally done by setting some thin membrane into vibration, for example by drawing a finely toothed leg or file across a wing panel, as in crickets, or by using a muscle to cause the progressive collapse of a stiff ribbed membrane covering a resonant body cavity, as in the case of cicadas. The result is generally a band of frequencies with little or no fine structure and a relative width of about 10 to 30 percent. The frequency scaling rule might be expected to be broadly similar for these cases, though perhaps displaced in reference frequency compared with vertebrates.

This discussion will not attempt to deal in detail with two other general cases. The first is scaling of the frequencies used by animals for echo location, for the optimization criteria are then very different from those for conspecific communication. The second is the case of aquatic animals, because the close match between the wave impedance of

a)Electronic mail: neville.fletcher@anu.edu.au

water and of body tissue introduces a new feature, and because propagation and attenuation in an ocean or lake environment, being nearly two-dimensional over long distances, has several features that are very different from an air environment.

II. SCALING ANALYSIS FOR VERTEBRATES

The lung pressure p driving the vocal air flow depends upon the pressure in the animal's lungs, which is in turn proportional to the thickness of the abdominal muscle walls and inversely proportional to the linear dimensions of the lung sac. These two variations cancel in animals of similar anatomy, so it can be concluded that vocalization pressure is approximately independent of animal size. The oscillating volume flow of air through the vocal valve is therefore

$$U=\left(\frac{2p}{\rho}\right)^{1/2}A_{\mathrm{v}}, \tag{1}$$

where A_{v} is the wide-open area of the valve and ρ is the density of air. If the oscillation frequency of the vocal valve is f, then since, in the cases we consider, the diameter of the valve and of the animal's mouth are both much less than the sound wavelength, the radiated sound power P can be shown[5,7] to be

$$P=\frac{\pi\rho f^2U^2}{2c}, \tag{2}$$

where c is the speed of sound in air. As the sound propagates, its intensity decreases, first because it is spread over a larger area, giving an inverse-square-law dependence, and second because of molecular atmospheric absorption.

Atmospheric absorption γ depends in a rather complex way upon frequency, temperature, pressure, and humidity, but overall it increases with frequency about as $\gamma=\alpha f^n$, with n close to 1.5 and α a constant with numerical value about $3.6\times10^{-8}\,\mathrm{m}^{-1}\,\mathrm{Hz}^{-1.5}$, in an atmosphere of normal humidity.[8,9] This gives a sound attenuation γ due to atmospheric absorption of about 0.5 dB/100m at 1 kHz, and correspondingly less at lower frequencies. The sound intensity $I(r)$ at a distance r from the source is therefore

$$I(r)\approx\frac{P}{4\pi r^2}\exp(-\alpha f^n r). \tag{3}$$

In more realistic evolutionary environments, such as forests of grasslands, a rather similar behavior can be expected, but with a higher value of the proportionality constant α.

The acoustic stimulus $S(r)$ provided to the auditory nerves of a listening animal of the same species at distance r is

$$S(r)=I(r)A_{\mathrm{e}}, \tag{4}$$

where A_{e} is the area of the external ear. Ignoring for the present the problem of interfering noise background, the sound will be audible at the distance r provided $S(r)$ exceeds some threshold value T that should not vary significantly from one species to another, if it is assumed that their neural transduction mechanisms are equally efficient. (This is, of course, an assumption, but there is no compelling evidence to suggest any systematic variation of efficiency with size. It is, of course, possible that some species have hearing abilities differing markedly from what is expected, and this may explain in part the scatter of observational data.) In summary then, the audibility requirement is that

$$S(r)\equiv\frac{pf^2A_{\mathrm{v}}^2A_{\mathrm{e}}}{4cr^2}\exp(-\alpha f^n r)\geqslant T. \tag{5}$$

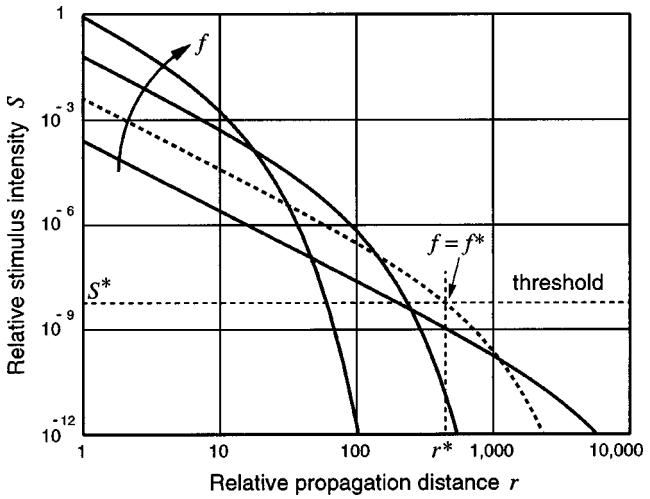


FIG. 1. Radiated sound intensity as a function of distance for a range of sound frequencies, assuming other parameters remain constant. The frequency increases in equal steps in the direction indicated. If the threshold level is as shown, then there is an optimal frequency f^* giving maximum range r^*, as shown by the broken curve.

The problem is now to vary the frequency f so as to maximize the distance r at which the animal call remains audible to members of the same species. The behavior of $S(r)$ as a function of r for various values of f is shown in Fig. 1. For high frequencies the sound level close to the animal is high, but it falls off rapidly with distance because of molecular absorption in the air, while at low frequencies the near-field sound level is lower, but it falls off less rapidly with distance. At the optimal frequency f^*, the range r^* at which the sound level equals the hearing threshold T is a maximum, so that the derivative $dr/df=0$, and this relation can be used in Eq. (5) to deduce that, at this optimum, $\alpha f^n r=2/n$. Substituting this back in (5) gives

$$f^*=\left(\frac{16cT}{\alpha^2n^2A_{\mathrm{v}}^2A_{\mathrm{e}}p}\right)^{1/(2n+2)}. \tag{6}$$

Assuming that both mouth and ear diameters are proportional to the average linear body dimension L gives the result that f^* is proportional to $L^{-3/(n+1)}$. Since the animal's mass M is proportional to L^3, this is equivalent to stating that the optimal communication frequency f^* is proportional to $M^{-1/(n+1)}$. Inserting the approximate value $n=1.5$ gives the result that f^* should be proportional to $M^{-0.4}$.

A. Experimental data

It is difficult to obtain good data from which to plot the way in which fundamental vocalization frequency depends upon body mass for a wide range of animals. This is partly

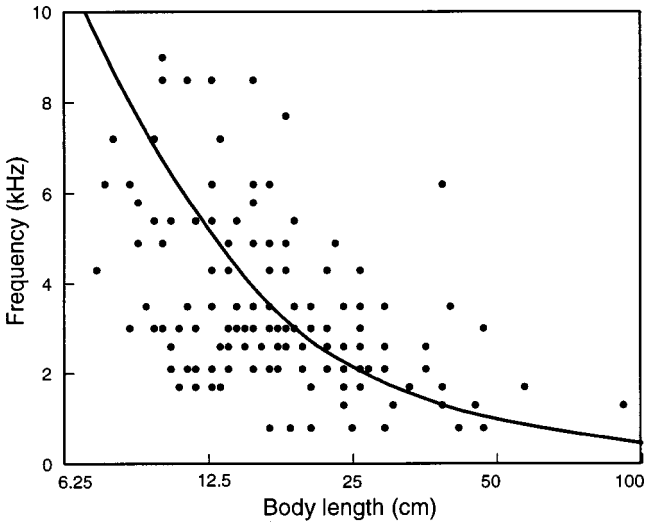


FIG. 2. Replot of the data of Ryan and Brenowitz (Ref. 10) on birdsong dominant frequency, with a superimposed curve of the form $f=AL^{-1.2}$ or $f=BM^{-0.4}$, as predicted by the theory given here.

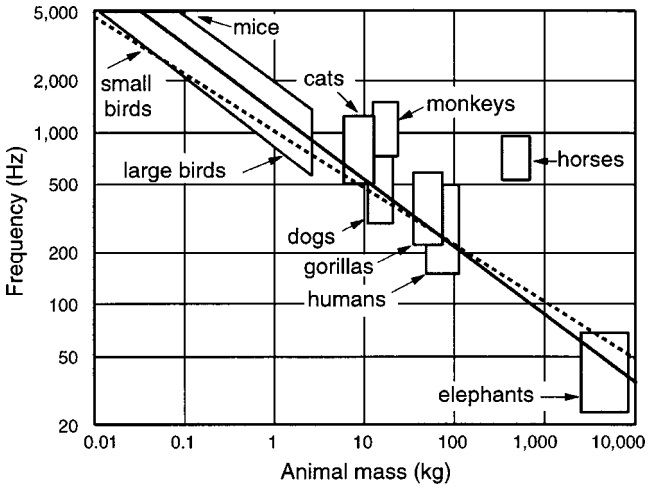


FIG. 3. Correlation between body mass and vocalization frequency for the animals shown. The full line shows the $M^{-0.4}$ power relationship predicted in the text and the dashed line the $M^{-0.33}$ relationship derived from simple linear size scaling.

due to the variability of the behavior of individual animals, and partly to variation between measurement techniques and decisions about exactly what is to be recorded. For the most part researchers have recorded the ''dominant frequency,'' which is probably a good basis for comparison.

Ryan and Brenowitz[10] report a very wide set of measurements on three different classes of birds—nonpasserines, suboscines, and oscines—in three different types of habitats. The data points for dominant frequency—the maximum in the spectrum—are rather widely scattered and were in each case fitted with an arbitrary regression of the form $f=A-B\log L$, where L is body length and A and B are positive constants. (This regression actually predicts a negative song frequency for birds larger than those plotted!) Figure 2 shows the complete data set. The regression line has been omitted since it is not meaningful. The complete data for birds are, however, fitted very well by a curve of the predicted form $f=BM^{-0.4}$, or equivalently $f=AL^{-1.2}$, as shown in Fig. 2, A being the only variable parameter. A regression curve of the form $f=CM^{-0.33}$, as conjectured from simple linear scaling would, however, fit the data just as well.

There is great difficulty in finding comparably extensive data for other groups of vertebrates, partly because of the wide variation in the sort of calls produced by higher animals. The published measurements, understandably, do not give the approximate body mass, so that it has been necessary to estimate this, and are mixed between fundamental frequency and dominant frequency. Figure 3 shows the approximate range of the dominant frequency of a wide set of animal vocalizations, plotted against their approximate body mass, using whatever data have been found. Data sets used, in part, to construct this figure have been published in individual chapters of the volume edited by Busnel[3] by Zhinkin, Dumortier, Bremond, and Tembrock, but other sources have also been used. Drawn in the figure is a line of the predicted slope -0.4 passing through the calculated optimal frequency for humans, and it is clear that this gives a reasonably good general account of the evolutionary optimization behavior over a range of about 10^6 in body mass. A considerable range of deviation from the theoretical line is to be expected, largely because of variation in the social patterns for which communication is used. Also shown in the figure is a broken line representing the simple linear scaling $f=M^{-0.33}$. This is seen to fit just as well.

B. Communication distances

Within a general class of species such as land mammals, or birds, the available lung pressure should be approximately independent of size, as discussed before. This then allows us to calculate the relative conspecific communication range r^* by setting $f=f^*=(2/n\alpha r)^{1/n}$, as was deduced from the differentiation of Eq. (5). The result is that r^* should be about proportional to $M^{0.6}$. No experimental data are immediately available to check this conclusion, but it does seem established that large mammals, such as elephants, can generally communicate over very much larger distances than small mammals, such as cats or mice, and large birds such as cockatoos over greater distances than small birds such as finches.

C. Actual frequencies

Although the scaling analysis outlined above is informative, the actual predicted scaling law for frequency is very close to the rule $f\propto M^{-1/3}$ that would be expected for a simple inverse scaling of frequency with body length. It is therefore important to go to the next stage and predict actual optimal communication frequencies. As a first step this can be done for the free-atmosphere case, though the effect of forest or grassland absorption should then be included. It must also be noted that, in most cases, vocal information is carried not just at the fundamental frequency but rather in the vocal formants, which have a frequency centroid at about 10 times the fundamental frequency in most animals. This sort of scaling is to be expected since the formants are produced

by longitudinal resonances of the vocal tract, the frequency of which varies inversely with this length and this with animal size.

Equation (6) requires some anatomical assumptions to lead to a predicted vocalization frequency, and an appropriate example is for a human when shouting loudly. In this case, it is a reasonable approximation to take $A_v \approx 10^{-5}\,\mathrm{m}^2$, $A_e \approx 10^{-3}\,\mathrm{m}^2$, and $p \approx 1000\,\mathrm{Pa}$. The normal auditory threshold in the middle of the vocalization band around 1000 Hz is about $10^{-12}\,\mathrm{W\,m}^{-2}$, so that $T \approx 10^{-15}\,\mathrm{W}$. Equation (6) then predicts an optimal vocalization frequency of about 450 Hz, which is about three times the fundamental frequency of human male vocalization, and about equal to that of female vocalization, and so approximately in the range of maximum vocal output. This is also close to the frequency of the first vocal-tract resonance. Given that the variation of vocalization fundamental frequency with body mass is quite well predicted, as indicated in Figs. 2 and 3, this means that there is probably comparable agreement in absolute predicted vocalization frequencies across the whole range of species and body masses encompassed in that plot.

As noted before, however, this optimal frequency varies with the air absorption αf^n as shown in (6). Assuming that $n=1.5$ in forest or grassland environments, and that the only change is an increase in the value of α, the prediction is that $f^* \propto \alpha^{-1/(n+1)} = \alpha^{-0.4}$, so that these environments would lead to a lowering of the optimal vocalization frequency. An increase in α by a factor 10 would lower the optimal frequency by a factor close to 0.4, giving 180 Hz for human vocalization. This is not such a good fit as the original free-air prediction. Such measurements as are available,[9] however, suggest that the variation of absorption with frequency is very much more complex in these environments than a simple power-law relation, so that this modification of the free-air prediction is perhaps not justified.

D. Background noise

There is, of course, another possible assumption that could be made about the optimization process, and that is that the criterion should be to maximize the signal-to-noise ratio at the receiving position, rather than just the signal level. In the modern world, a large proportion of background noise in places of human habitation is due to human activities such as transport, but this is too recent to have had any evolutionary impact. Background noise due to other causes, predominantly air turbulence, ocean waves, and biological sounds, has a frequency distribution that varies greatly with local environment, but on a larger scale has about constant noise power in each octave band across the spectrum. Since the auditory systems of vertebrate animals have roughly a constant fractional bandwidth $f_{\max}/f_{\min}$ of about 4 octaves (a factor 16) at the 10-dB point, this means that they tend to receive a constant noise input, regardless of their center frequency.

Assuming that the external ear is large enough that noise is, in fact, the limiting factor, the optimization of communication frequency therefore amounts to maximizing $I(r)$ rather than $I(r)A_e$ for each size of animal. Going through the optimization procedure as before leads to the scaling law $f^* \propto M^{-2/(3n+3)}$, or $f^* \propto M^{-0.27}$. While not completely ruled out by the available data, this power law appears to fit it rather less well. It is possible, however, that the animal neural system can disregard noise lying outside the immediate vicinity of the signal frequency, a situation known as selective masking, in which case a rather different conclusion can be reached,[5] with the optimal frequency depending upon the available signal power.

Of course, many animals live in groups in noisy environments such as rainforest, and the optimization problem for them may be rather different. It hardly seems possible to decide much on this point except by looking at the optimization while taking into account the particulars of the animal and its situation, and assuming these have not varied greatly over the evolutionary time involved. This doubtless accounts for some of the variation observed between different species.

III. INSECTS

A similar analysis could be carried out for insects that produce sound by stridulation—the setting of wing panels into vibration through the action of a file on their leg. In this case, because the panel has two free sides, the source is an acoustic dipole, rather than a monopole, and this inserts an extra factor f/r into $I(r)$, but, in addition, the source amplitude now varies as d^3, rather d^2, where d is the diameter of the vibrating panel, because the effective distance between the two poles of the dipole is also proportional to d. Following through the argument as before, we conclude that the optimal frequency varies as $M^{-0.36}$ and the conspecific communication range as $M^{0.53}$ within a class of sound-producing insects with broadly similar anatomy. These scaling laws are broadly similar to those applying to mammals, but the constant of proportionality will, however, be different. Because insects are so much smaller than vertebrates, their communication frequencies are generally much higher, typically 3 to 5 kHz, in accord with the general trend of these scaling laws.

While most mammals devote a small and nearly constant fraction of their available energy to sound production, the situation is often very different in insects. Some Australian cicadas, for example, produce a sound power of around 1 mW at about 3 kHz, which is about equal to that produced in human singing. In another variety of Australian cicada of about the same length, *Cystosoma Saundersii* (Westwood), the male anatomy is so specialized for sound production[11] that the abdomen, which makes up nearly 90 percent of the volume of the animal, is a simple hollow sac that reinforces the vibration of the tymbals at the low-song frequency of 800 Hz. Some ground-dwelling crickets even dig burrows in the shape of exponential horns, so that an animal of mass less than 1 gram has an effective acoustic mouth diameter of several centimeters for sound radiation.[12]

IV. AQUATIC ANIMALS

In the case of aquatic mammals, similar evolutionary pressures should also have applied. Sea water has a similar variation of sound absorption with frequency to that of air,[13] but the sound absorption in sea water is much less than that in air, leading to vastly increased communication distances.

Over such distances, which are much greater than the ocean depth, sound propagation is essentially two-dimensional, which adds further to the propagation distance. In addition to this, aquatic mammals do not simply discharge an oscillating airflow into the surrounding water, but rather generally discharge it from one body cavity to another. The linear dimensions of these cavities, however, are generally proportional to animal length, so the scaling is preserved. Carrying through the optimization in the same way leads to the conclusion that $f^* \propto M^{-3/(n+2)}$ or approximately $f^* \propto M^{-0.6}$, and thus a much greater predicted variation with size than for land mammals. Optimization has, however, taken a different direction with those mammals that use sound primarily for sonar detection rather than for conspecific communication.

In the case of sound-producing crustaceans, the scaling law should be very much like that for insects, perhaps modified for two-dimensional propagation since these animals often live in shallow water. This gives $f^* \propto M^{-0.44}$ for shallow water and $f^* \propto M^{-0.6}$ for deep water. Since, however, the clicking sounds produced by these animals seem to be by-products of their activity rather than optimized for long-distance communication, their frequencies may indeed be simply related to body size as $M^{-0.33}$. Variation of body shape between species makes it unlikely that any general rule applies.

V. CONCLUSION

While it is possible that vocal communication frequencies in animals are not optimized at all, but are simply the result of overall blind size scaling of the vocal apparatus which would give a power law $f^* \propto M^{-0.33}$, it seems more likely that the sort of evolutionary optimization discussed here has actually occurred, giving an approximate power law $f^* \propto M^{-0.4}$. Because these two predicted rules do not vary greatly in exponent, it is probably not possible to make a decision on this point, but it is interesting that the sort of optimization that would be predicted, based upon simple physical arguments, is so close to that found in nature.

Despite all the variations in purpose and method of conspecific communication, it is, we believe, instructive to examine general trends in the way that has been done here. Similar analyses could be made of the generation and detection of vibrations by insects that use this technique for communication and prey detection, and of the sound emissions by animals such as bats and dolphins, the purpose of which is echolocation rather than conspecific communication.

[1] W. C. Stebbins, *The Acoustic Sense of Animals* (Harvard University Press, Cambridge, MA, 1983).
[2] J. W. Bradbury and S. L. Vehrenkamp, *Principles of Animal Communication* (Sinauer Sunderland, MA, 1998).
[3] *Acoustic Behavior of Animals*, edited by R.-G. Busnel (Elsevier, New York, 1963).
[4] *Bioacoustics: A Comparative Approach*, edited by B. Lewis (Academic, London, 1983).
[5] N. H. Fletcher, *Acoustic Systems in Biology* (Oxford University Press, New York, 1992).
[6] N. H. Fletcher and S. Thwaites, ''Physical models for the analysis of acoustical systems in biology,'' Q. Rev. Biophys. **12**, 26–66 (1979); **12**, 463 (1979).
[7] P. M. Morse, *Vibration and Sound* [Acoustical Society of America, New York (reprinted 1984)], Chap. 7.
[8] H. E. Bass, L. C. Sutherland, A. J. Zuckerwar, D. T. Blackstock, and D. M. Hester, ''Atmospheric absorption of sound: Recent developments,'' J. Acoust. Soc. Am. **97**, 680–683 (1995).
[9] L. C. Sutherland and G. A. Daigle, ''Atmospheric sound propagation,'' in *Encyclopedia of Acoustics*, edited by M. J. Crocker (Wiley, New York, 1997), pp. 341–365.
[10] M. J. Ryan and E. A. Brenowitz, ''The role of body size, phylogeny, and ambient noise in the evolution of bird song,'' Am. Nat. **126**, 87–100 (1985).
[11] N. H. Fletcher and K. G. Hill, ''Acoustics of sound production and hearing in the bladder cicada *Cystosoma Saundersii* (Westwood),'' J. Exp. Biol. **72**, 43–55 (1978).
[12] A. G. Daws, H. C. Bennet-Clark, and N. H. Fletcher, ''The mechanism of tuning of the mole cricket singing burrow,'' Bioacoustics **7**, 81–117 (1996).
[13] R. E. Francois and G. R. Garrison, ''Sound absorption based on ocean measurements. II. Boric acid contribution and equation for total absorption,'' J. Acoust. Soc. Am. **72**, 1879–1890 (1982).

Localization of aerial broadband noise by pinnipeds

Marla M. Holt, Ronald J. Schusterman, Brandon L. Southall, and David Kastak
Long Marine Laboratory, University of California, Santa Cruz, 100 Shaffer Road, Santa Cruz, California 95060

(Received 18 August 2003; revised 30 January 2004; accepted 10 February 2004)

Although many pinnipeds (seals, sea lions, and walruses) emit broadband calls on land as part of their communication system, few studies have addressed these animals' ability to localize aerial broadband sounds. In this study, the aerial sound localization acuities of a female northern elephant seal (*Mirounga angustirostris*), a male harbor seal (*Phoca vitulina*), and a female California sea lion (*Zalophus californianus*) were measured in the horizontal plane. The stimulus was broadband white noise that was band pass filtered between 1.2 and 15 kHz. Testing was conducted in a hemi-anechoic chamber using a left/right forced choice procedure to measure the minimum audible angle (MAA) for each subject. MAAs were defined as half the angular separation of two sound sources bisected by a subject's midline that corresponded to 75% correct discrimination. MAAs were 4.7°, 3.6°, and 4.2° for the northern elephant seal, harbor seal, and California sea lion, respectively. These results demonstrate that individuals of these pinniped species have sound localization abilities comparable to the domestic cat and rhesus macaque. The acuity differences between our subjects were small and not predicted by head size. These results likely reflect the relatively acute general abilities of pinnipeds to localize aerial broadband signals.
[DOI: 10.1121/1.1694995]

PACS numbers: 43.80.Lb, 43.66.Qp [WAU] Pages: 2339–2345

I. INTRODUCTION

Pinnipeds are a group of carnivorous mammals that comprise three families: Phocidae (true seals), Otariidae (sea lions and fur seals), and Odobenidae (walruses). These amphibious animals typically forage and navigate under water but are tied to land or ice for molting and at least some components of their reproductive cycle. Pinnipeds rely on a variety of acoustic signals in both media to coordinate important life history events. In general, pinniped airborne calls are stereotyped and broadband (Bartholomew and Collias, 1962; Peterson and Bartholomew, 1969; Perry and Renouf, 1988) and contain directional components that are dependent on the orientation of the caller (Schusterman, 1978; Southall, 2002). Sound emissions are largely used for communicative purposes particularly in the context of reproductive activities. In general, calls are produced under water in odobenids and aquatically breeding phocids and in air in otariids and land or ice breeding phocids. For example, mutual calling in air functions in the maintenance of contact between mothers and offspring during the lactation period in otariids (Trillmich 1981; Gisiner and Schusterman, 1991; Insley, 2001) and elephant seals (Bartholomew and Collias, 1962; Petrinovich, 1974). Furthermore, studies have shown that otariid mothers and pups are capable of recognizing each others calls on land (e.g., Insley 2001; Charrier *et al.*, 2002, 2003). Males produce aerial vocalizations while delineating territorial or hierarchical status during the breeding season (Bartholomew and Collias, 1962; Peterson and Bartholomew, 1969). In addition to intraspecific signaling, aerial sounds may also be important for detecting and avoiding land predators [e.g., bears and foxes (Stirling and Archibald, 1977)]. Thus, sounds serve important biological functions in pinnipeds and the ability to determine the accurate location of these sounds can have significant consequences on an individual's fitness.

Several investigations on hearing in marine mammals have recently been conducted, in part due to concerns about increased levels of noise in the ocean (Richardson *et al.*, 1995; Andrew *et al.*, 2002). These studies describe aerial hearing abilities of several pinniped species and demonstrate that hearing in some species can be affected by exposure to noise (Turnbull and Terhune, 1990; Kastak and Schusterman, 1996, 1998; Kastak *et al.*, 1999; Southall *et al.*, 2000, 2003). While these studies have contributed to a larger understanding of the peripheral mechanisms involved in pinniped hearing, little is known about other auditory abilities of pinnipeds such as sound localization. Sound localization capabilities have been measured in only three pinniped species: northern fur seals in air and under water (Babushina, 1998; Babushina and Yurkevich, 1994), two California sea lions under water (Gentry, 1967; Moore, 1975; Moore and Au, 1975), and two harbor seals in air and under water (Møhl, 1964; Terhune, 1974). A comparative assessment of pinniped aerial and underwater sound localization revealed that localizing broadband sounds under water is more difficult for harbor seal subjects (Terhune, 1974). Despite the social importance of localizing broadband signals in air by all land breeding pinnipeds, only two studies (Terhune, 1974; Babushina, 1998) have measured this ability in the horizontal plane in these animals.

The typical mammalian auditory system primarily utilizes two types of binaural cues for sound localization in the horizontal plane: interaural phase differences and interaural level differences (Yost, 2000). Interaural phase differences (IPDs) are generally important for localizing low frequency sounds given phase ambiguity at higher frequencies. Interau-

ral level differences (ILDs), primarily created by the shadowing effects of the head, are generally important for localizing high frequency sounds (for review, see Heffner and Heffner, 1992a). The dual processing of these two binaural cues describes the duplex theory of sound localization (Strutt, 1907) and has been supported by several studies conducted in humans (Mills, 1958; Wightman and Kistler, 1992; Macpherson and Middlebrook, 2002) and in other mammals (Heffner and Heffner, 1992a) including a harbor seal in air and a California sea lion under water (Moore and Au, 1975; Terhune, 1974).

The most common metric for measuring sound localization ability is the minimum audible angle (MAA). The MAA is most often defined as half the angle between two identical sound sources bisected by a subject's midline that are acoustically discriminated as separate in space (Mills, 1958), with smaller MAAs representing better localization ability. Because the ability to localize sound in the horizontal plane is generally dependent on both time and intensity based binaural cues, testing is often conducted using broadband signals that contain both IPDs and ILDs. In mammals, horizontal plane MAAs for broadband signals range from a few degrees in humans (Mills, 1958), elephants (Heffner and Heffner, 1992a), and dolphins (under water, Renaud and Popper, 1975) to over 30° in cows (Heffner and Heffner, 1992b) and gophers (Heffner and Heffner, 1990). Because binaural cues are theoretically enhanced by a larger head, some investigators have hypothesized a negative correlation between MAA and head size in mammals. While some of the available data suggest such a correlation, there are numerous exceptions. For instance, cows have relatively large heads but are poor sound localizers (Heffner and Heffner, 1992b). Thus, head size alone does not comprehensively account for the variations of sound localization abilities measured in mammals tested thus far (see Heffner and Heffner, 1992a). Furthermore, not all mammals have demonstrated the ability to utilize both binaural cues for horizontal plane sound localization. For example, evidence from behavioral testing indicates that hedgehogs have a limited ability to utilize IPDs (Masterton *et al.*, 1975) while horses have a limited ability to utilize ILDs (Heffner and Heffner, 1986).

In this study, we determined aerial MAAs in the horizontal plane using a broadband white noise signal in trained individuals representing three pinniped species: a northern elephant seal (*Mirounga angustirostris*), a harbor seal (*Phoca vitulina*), and a California sea lion (*Zalophus californianus*). The elephant seal and California sea lion are land breeders and primarily call in air and the harbor seal is an aquatic breeder and primarily calls under water. This allowed us to investigate the overall sound localization abilities of these pinnipeds given the possible ecological and physical factors that may influence this auditory ability (e.g., head size).

II. METHODS

A. Subjects

The pinniped subjects were a 9-year-old female northern elephant seal (Burnyce), a 13-year-old male harbor seal (Sprouts), and a 16-year-old female California sea lion (Rio). All subjects were kept at Long Marine Laboratory in Santa Cruz, CA in free-flowing saltwater pools with adjacent haul-out sites. All subjects had previous experience performing sound detection tasks in air and under water and appeared to have normal hearing (Kastak and Schusterman, 1998; Kastak *et al.*, 1999; Southall *et al.*, 2000). In the current study, subjects received approximately 30%–50% of their daily food total (3–15 kg of mixed herring and capelin) during experimental sessions. Additionally, one 22-year-old female human subject (KAJ) was tested under identical conditions so that her results could be compared with previously published data on human subjects. Prior to localization testing, this subject was tested to ensure that her absolute hearing sensitivity was typical for her age group (ANSI 3.6-1996). This study followed the protocols approved by the University of California Chancellor's Animal Research Committee (CARC) and the University of California Human Subjects Institutional Review Board (IRB), and was conducted under National Marine Fisheries Services (NMFS) permit #259-1481-00.

B. Apparatus

Testing occurred in a custom-built $4.0\times2.8\times2.4$-m double-walled hemi-anechoic chamber (Eckel Industries). The surfaces were lined with acoustic foam wedges except the concrete floor, which was covered with 2.6-cm-thick neoprene mats, and the stainless steel door in the rear of the chamber. The testing chamber was lit with two 75-W light bulbs and was ventilated with a remote quieted fan during testing. A stationing chin cup made of PVC was mounted to the floor and placed 1 m from the sound sources mounted to the wall in front of the chin cup. Two rectangular PVC response targets (each 11.4×8.9 cm) were also mounted to the floor and spaced equi-distance (11.7 cm) to either side of the chin cup. All subjects were tested with stimuli projected at ear level. Given the height differences between the subjects, the height of the chin cup was 33.0 cm for the harbor seal and sea lion and 38.1 cm for the elephant seal. The height of the response targets was 33.0 cm for all subjects. A small light was placed 1 m in front of the subjects and was manually turned on and off during each trial by the experimenter. This was done to ensure that the subject was attentive to the auditory signal during the trial interval.

C. Stimulus production and equipment

All equipment was controlled by the experimenter in a $1.3\times2.8\times2.4$-m^3, control room adjacent to the test chamber. The signal was a single noise burst that was generated and triggered by the experimenter using custom designed LabVIEW™ 6i software and operated by a National Instruments PXI 1010 Chassis and National Instruments 6070E multifunction I/O board. Unrepeated samples of white noise were produced on each trial with a rise and fall time of 0.1 ms to avoid switching transients when the signal was triggered. The signal was 100 ms in duration, which was short enough to minimize any performance improvement due to head motion during its presentation (Heffner *et al.*, 2001). The signal was band pass filtered between 1.2 and 15 kHz and thus included frequencies that spanned most of the aerial hearing

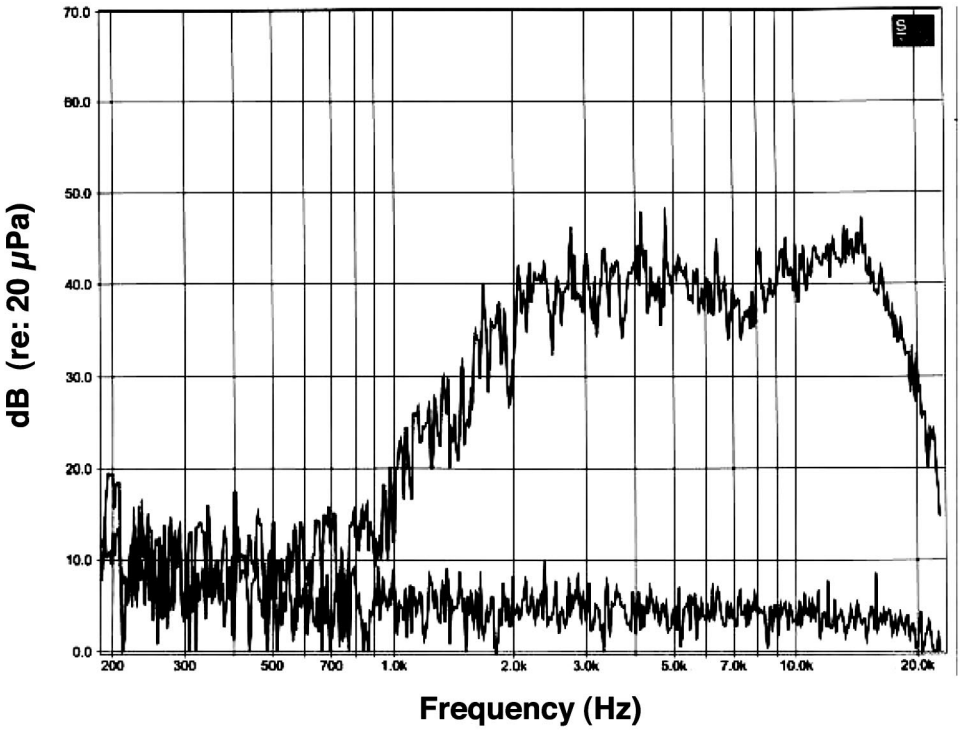


FIG. 1. Spectrum of the noise stimulus used in this study (based on 1.5-Hz analysis bandwidth). The top curve represents the stimulus and the bottom curve represents the position of the noise floor relative to the stimulus.

range of all subjects (Kastak and Schusterman, 1998). Figure 1 shows a plot of the noise spectrum (at 0° azimuth) used during experimental testing. The signal was routed to a speaker selector (Acoustic Research 1108) that allowed one of six connected speakers (Morel MDT37 horn tweeters) to be manually selected by the experimenter. Before each trial, two speakers were always selected, one from the left and one from the right and then one of these two speakers was made inactive before the beginning of a trial. This was done to prevent possible cuing from any switching transients produced during the intertrial period. Inside the test chamber, each speaker was housed in a PVC cup and mounted on a 2.4-m linear track that was placed 1 m away from and at 0° elevation relative to the subjects' ears. This allowed the speakers to be placed anywhere between 45° to the left and 45° to the right of the subjects' midline (with $\pm 0.5°$ of accuracy). Within a session, three speakers were always placed to the right and three speakers were always placed to the left of the subject's midline. The speaker positions were manually changed in azimuth relative to the subject's midline between experimental sessions. A modified method of constant stimuli was used to determine horizontal plane MAAs for each subject.

The average sound pressure level of the stimuli (dB *re*: 20 μPa) was determined at a position corresponding to the center of the subject's head (with the subject removed) using a calibrated microphone (Etymotic Research ER-7C) and a spectrum analyzer (SpectraPlus®) using narrow-band analysis (48 000 Hz sample rate, 32 768 FFT size, 1.5-Hz analysis bandwidth) for each of the test angles at the beginning of each experimental session. Acoustic mapping was performed prior to testing in which received noise stimulus levels were measured at every 2° within 10° of azimuth and at every 5° between 10° and 45° of azimuth. For each azimuth location, received levels were measured at ten separate positions within a $10\times10\times10$-cm^3 area surrounding the chin cup of the test apparatus. Differences of up to 10 dB in sound pressure levels were observed between stimuli projected from different speaker locations during acoustic mapping. Therefore, the stimulus level was randomly varied on each trial over a 12-dB range surrounding the average level (in 2-dB steps) so that the subjects could not use intensity discrimination between speakers or speaker positions to improve performance. Plots of the noise spectrum for both left and right angles at 2°, 4°, 6°, 8°, and 10° were compared against each other and at 0° using the overlay function on SpectraPlus® and showed that the spectra measured for each location were essentially identical, especially between symmetrical angles. Small differences between spectra likely resulted from generating new samples of noise on each trial (i.e., the signal was not a fixed sample of white noise that was repeated over and over again). Additionally, the position of any one speaker was periodically switched between sessions relative to other speakers in the array so that any differential performance between speakers was not systematically correlated with a position in the horizontal plane.

Sound detection thresholds for the signal were determined prior to localization testing for each subject using methods described in Kastak and Schusterman (1998). The average signal level was 60 dB above threshold for the California sea lion, harbor seal and human, and 35 dB above threshold for the northern elephant seal. The perceived signal level was lower for the elephant seal due to a lower sensitivity to the stimulus and limitations of the equipment. Based on the results of studies involving terrestrial mammals, it has been shown that differences in sensation level had little or no effect on the ability of the subjects to localize the signal (e.g., see Heffner *et al.*, 2001).

D. Procedure

A left/right forced choice procedure was used to determine the MAAs for each subject. A correct response was defined as pressing the left target when the left speaker was activated and pressing the right target when the right speaker was activated. This procedure was novel to all subjects of this study and each had extensive prior experience performing sound detection tasks using a go/no-go procedure (Southall *et al.*, 2003). A spontaneous orienting response to a nearby sound source was used to train the left/right discrimination procedure for this investigation. For the first phase of training, the two response targets were placed directly next to either side of the subject with a speaker located directly behind each target. Training consisted of approximating a left/right response by rewarding the subject for orienting towards, then approaching and then eventually nose-touching the correct target when the speaker played from behind it. In this phase of training, all three subjects showed a spontaneous orienting and then responding to targets associated with spatially disparate speakers. Then the speakers were moved successively away from the targets and towards the speaker track until eventually the responses were under good stimulus control using test speakers placed at $\leqslant 45°$ relative to the subject's midline on the track. All subjects achieved performance criterion of $\geqslant 90\%$ correct responses at angles wider than 15° by the ninth training session (each approximately 25 trials per session).

During data collection, subject responses were monitored by the experimenter in the control room via a surveillance camera. A trial began when the animal stationed properly in the chin cup and the trial light was manually illuminated by the experimenter. If a response was made to the left target after a left sound presentation or to the right target after a right sound presentation, a digitized whistle bridge was played through a separate speaker (placed on 0° azimuth) indicating to the subject that a correct response was made, the trial light was turned off, and an assistant sitting in the control room delivered a fish reward via a PVC conduit in front of the subject. Both types of correct responses were given an equal proportion of fish (i.e., a payoff matrix of 1:1). Additionally, one or two no-go trials in which the animal was reinforced for not responding to ''blank'' trials were randomly incorporated within each experimental session. Responses made before the stimulus presentation or to an incorrect target were not reinforced and the trial light was extinguished by the experimenter, indicating the end of the trial. The overall and first order conditional probabilities of left and right trial presentations were 0.5 within a testing session (Holt and Schusterman, 2002). Approximately six warm-up and six cool-down trials were given at the beginning and end of each experimental session, respectively, in which the stimulus was projected at 25° to the right and the left of the subject's midline. Warm-up, cool-down, and no-go trials were incorporated to ensure good stimulus control over the subject's behavior and were not used in the analysis to calculate MAAs.

E. Statistical analysis

After the initial training phase for each subject was completed, 25 trials at each test azimuth (between 1° and 12°, in 1° increments) were collected in random order in approximately 30 experimental sessions. The resulting data were pooled across all sessions for each subject and the percentages of right target responses were plotted against both left and right test angles (Mills, 1958). Performance for left and right test angles was plotted separately rather than pooling performance across test angles of symmetry to determine if any of the subjects had measurable biases to one side. All of the data from the resulting plots followed a sigmoidal psychometric function. Finney's (1971) probit analysis was used to linearize the data and interpolate MAAs from a linear regression analysis as the azimuth value corresponding to 75% correct performance for both left and right angles. Because performance for left and right angles were plotted separately, this average corresponds to the 75% right target responses for the right sound presentations (positive angles on the x-axis) and the 25% right target responses for the left sound presentations (negative angles on the x-axis) since 100 minus percent right target response on this plot corresponds to the percent left target response and thus correct performance for left angles. Probit analysis was used as opposed to assuming a purely linear model for the raw data because the decreasing slope of the psychometric functions at larger angles would otherwise inflate the estimated MAAs [for a brief discussion see Mills (1958)].

III. RESULTS

Subjects sometimes developed a bias towards one response within an experimental session, especially with smaller test angles (i.e., when discrimination was difficult). However, any biases observed within an experimental session were averaged out between sessions because none of the subjects developed long-term biases. Figure 2 shows the psychometric function for each subject plotted as the percentage of right target responses against both left and right test angles. Note that the proportion of left target responses is one minus the proportion for right target responses. Minimum audible angles, head radii and maximum interaural distance for each subject are shown in Table I. The maximum interaural distance was calculated for each subject as the time in microseconds it would take airborne sound to travel from one side of the head to the other using head diameter as the distance and the speed of sound in air equal to 343 m/s. Of the three pinniped subjects of this study, the harbor seal had the smallest MAA (3.6°, s.d.$=0.12$), followed by the California sea lion (4.2°, s.d.$=0.10$) and then the northern elephant seal (4.7°, s.d.$=0.11$), although the differences between subjects were small. Interpreting interspecific differences is problematic given the small sample size and likelihood that interindividual differences exist within a species.

Figure 3 shows the relationship between sound localization acuity in the horizontal plane and interaural distance for the pinnipeds that we tested in this study as well as other selected terrestrial mammals which are shown for comparison.

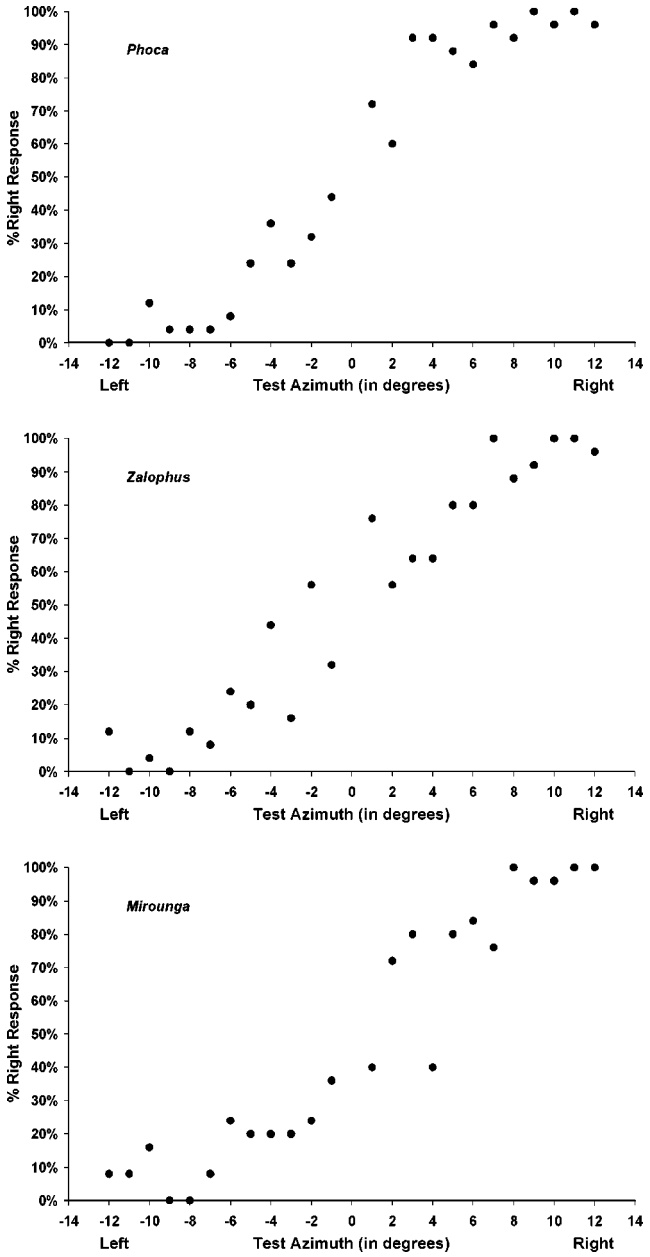


FIG. 2. Psychometric function for each animal tested in this study. For test angles presented to the right of the subjects' midline, correct responses correspond directly to the percent right target responses while for the left angles (negative angles on the *x*-axis), correct responses are 100 minus the percent right target response on the plot. Minimum audible angles were estimated by averaging the left and right test angles corresponding to 75% correct performance.

IV. DISCUSSION

In this study, MAAs were determined using a left/right forced choice procedure and the subjects sometimes developed response biases to one side within an experimental session. This bias was noted from differential performances between left and right azimuths of symmetry. This type of bias was also noted in previous studies with harbor seal subjects (Møhl, 1964; Terhune, 1974). Despite this caveat, a left/right procedure was used in this study for two reasons. First, the spontaneous orienting response of the subject to a sound source could be utilized for efficient training of the task. Second, all subjects had performed sound detection tasks using a go/no-go procedure prior to the sound localization task and training a left/right procedure seemed to enable the subjects to differentiate the two tasks more easily. Because the stimulus was presented on every trial, the subjects rarely made responses before the sound presentation. MAAs were statistically determined by plotting the proportion of right target responses against both left and right test azimuths in this study [as in Mills (1958)]. Investigators of other localization studies have taken a variety of different procedural and statistical approaches and thus methodological differences between this study and others should be carefully considered while making data comparison across studies.

As shown in Table I, the 0.6° MAA measured in the human subject is consistent with previously reported broadband MAAs in humans (Heffner and Heffner, 1992b). This suggests that there was nothing experimentally unusual in the testing conditions of this study. Furthermore, the human subject was interviewed after the completion of data collection and reported that no switching transients or background noise was detectable from any speaker during testing. The harbor seal MAA of this study closely matches those previously reported broadband MAAs measured in air for the same species [3° for click train stimulus, 75% correct for a left/right procedure; 2° for broadband noise, 50% correct for a center/left procedure, (Terhune, 1974)].

Head size alone did not account for the small MAA differences between the subjects of this study given that the subject with the largest head size (the northern elephant seal) had the largest MAA as shown in Table I. As can be seen from Fig. 3, some mammals with relatively large interaural distances are generally poor sound localizers (e.g., cows and goats). Heffner and Heffner (1992b) have addressed a number of other variables such as predation, domestication, and vision as possible correlates of sound localization acuities measured in mammals. They have shown that sound localization acuity is inversely correlated with the width of the visual field such that mammals with broad visual fields have relatively poor sound localization acuities and vice versa (Heffner and Heffner, 1992a; Heffner *et al.*, 2001). This supports the hypothesis that one of the main roles of sound localization in terrestrial mammals is to orient the subject to the sound source for visual purposes (Heffner and Heffner, 1992a). Such a spontaneous orienting response to a nearby sound source was observed in all three pinnipeds of this study and utilized for training the left/right procedure. Furthermore, such an orienting response not only directs the visual field of the subject toward the source but it also positions the subject so that sound is projected relatively closer to the midline. This may be advantageous, particularly in low-light conditions because localization acuity in the horizontal plane, at least in humans, is best at the midline compared to acuity measured with sources projected at wider angles (e.g., Mills, 1958). With the exceptions of walruses, pinnipeds have eyes directed forward and probably have good binocular vision (Riedman, 1990). However, quantitative measurements of the widths of the visual field for the species of this study are not currently available. Such information would allow us to determine if this variable predicts the MAAs measured in this investigation.

TABLE I. Aerial minimum audible angles (in degrees azimuth), standard deviations, head radii (in meters), and interaural distance (in microseconds) for each subject of this study.

Subject	MAA (degrees)	s.d.	Head radius (m)	Interaural distance (in μsec)
Phoca vitulina (Sprouts) 13-year-old male	3.6	0.12	0.083	760
Zalophus californianus (Rio) 16-year-old female	4.2	0.10	0.080	733
Mirounga angustirostris (Burnyce) 9-year-old female	4.7	0.11	0.154	1411
Human subject (KAJ) 22-year-old-female	0.6	0.01	0.090	824

Given the importance of localizing conspecific calls on land and the small MAA differences between the subjects of this study, horizontal plane sound localization of broadband signals in most pinniped species is probably relatively acute. This study illustrates the aerial localization ability of three pinnipeds including two phocid and one sea lion species. Of the two phocid subjects of this study, one is a land breeder and primarily calls in air (the Northern elephant seal) while the other is an aquatic breeder and primarily calls under water (the harbor seal). The sound localization abilities of these two subjects were not considerably different despite environmental differences in their vocal reproductive behavior. Of the mammalian data available, broadband sound localization abilities for all pinniped subjects of this study are comparable to pigs, opossums, cats, and rhesus macaques with only humans, elephants, and dolphins having better acuity (Renaud and Popper, 1975; Heffner and Heffner, 1992a). It is very likely that the pulsed, broadband, directional and repetitive nature of pinniped calls in air aids in both the detection and localization of these sounds in naturally noisy environments (Schusterman *et al.*, 2000). A few studies have measured localization of a broadband signal under water in a California sea lion and a harbor seal and results show that localization is more difficult in this medium than in air (Terhune, 1974; Moore, 1975). This is an important finding considering the amphibious nature of pinnipeds and suggests that the faster travel time of underwater sounds may constrain this auditory ability. Other sensory modalities such as

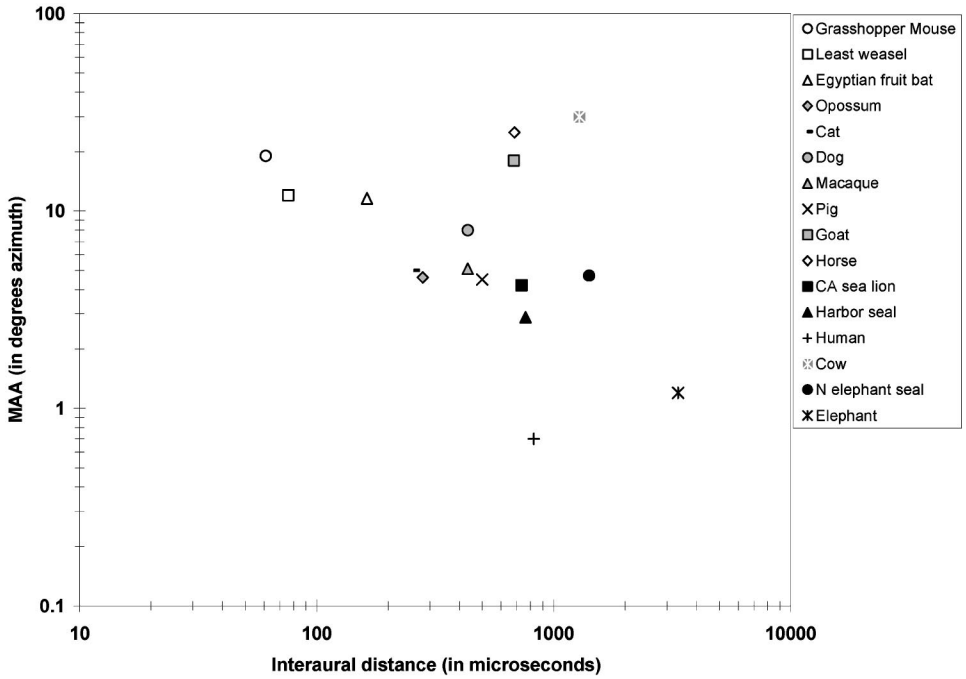


FIG. 3. Aerial minimum audible angles of broadband stimuli plotted against interaural distance (in microseconds) for selected land mammals and pinnipeds. Legend lists species in increasing order of interaural distance (note log-log scale). Grasshopper mouse, $n=2$ (Heffner and Heffner, 1992a); least weasel, $n=3$ (Heffner and Heffner, 1992a); Egyptian fruit bat, $n=3$ (Heffner *et al.*, 2001); cat, $n=5$ (Heffner and Heffner, 1992a); opossum, $n=6$ (Ravizza and Masterton, 1972); dog, $n=4$ (Heffner and Heffner, 1992b); macaque, $n=6$ (Brown *et al.*, 1980); pig, $n=3$ (Heffner and Heffner, 1992a); goat, $n=2$ (Heffner and Heffner, 1992b); horse, $n=2$ (Heffner and Heffner, 1992a); California sea lion, $n=1$ (this study); harbor seal, $n=2$ (this study; Terhune, 1974); human, $n=4$ (this study, Heffner and Heffner, 1992b); cow, $n=3$ (Heffner and Heffner, 1992b); northern elephant seal, $n=1$ (this study); and elephant, $n=1$ (Heffner and Heffner, 1992a).

vision may be more important for underwater activities involving traveling and foraging (see Schusterman *et al.*, 2000). Additional localization investigations involving pure tone and underwater signals would provide further insight regarding the use of binaural cues and differences in localization abilities in air versus under water in these amphibious animals.

ACKNOWLEDGMENTS

This study was made possible by funding from the Office of Naval Research (Grant No. N00014-99-1064) to RJS and the Myers' Oceanographic Trust Fund, the Friends of Long Marine Laboratory, and the Ocean Sciences Department at the University of California, Santa Cruz to MMH. The authors would like to thank Colleen Reichmuth Kastak for her insight and assistance on animal training and procedure and for her thoughtful comments on earlier versions of this manuscript. Special thanks to the human subject, KAJ, for her participation in this study and the volunteers at Long Marine Laboratory, Santa Cruz, CA for their continual support of our research in pinniped cognition and sensory systems.

Andrew, R. K., Howe, B. M., Mercer, J. A., and Dzieciuch, M. A. (**2002**). ''Ocean ambient sound: Comparing the 1960s with the 1990s receiver off the California coast,'' ARLO **3**, 65–70.

ANSI (**1996**). ANSI-3.6-1996, ''Specifications for audiometers'' (American National Standards Institute, New York).

Babushina, E. S. (**1998**). ''Localization of an acoustic signal source in the vertical plane in air by the northern marine seal (*Callorhinus ursinus*),'' translated from Sen. Sist. **12**, 444–451.

Babushina, E. S., and Yurkevich, L. I. (**1994**). ''Vertical plane localization of underwater sound sources by the northern fur seal,'' Translated from Sens. Sist. **8**, 55–57.

Bartholomew, G. A., and Collias, N. E. (**1962**). ''The role of vocalization in the social behavior of the northern elephant seal,'' Anim. Behav. **10**, 7–14.

Brown, C. H., Beecher, M. D., Moody, D. B., and Stebbins, W. C. (**1980**). ''Localization of noise bands by old world monkeys,'' J. Acoust. Soc. Am. **68**, 127–132.

Charrier, I., Mathevon, N., and Jouventin, P. (**2002**). ''How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*.'' J. Exp. Biol. **205**, 603–612.

Charrier, I., Mathevon, N., and Jouventin, P. (**2003**). ''Vocal signature recognition of mothers by fur seal pups'' Anim. Behav. **65**, 543–550.

Finney, D. J. (**1971**). *Probit Analysis*, 3rd ed. (Cambridge U.P., Cambridge).

Gentry, R. L. (**1967**). ''Underwater auditory localization in the California sea lion (*Zalophus californianus*),'' J. Aud. Res. **7**, 187–193.

Gisiner, R., and Schusterman, R. J. (**1991**). ''California sea lion pups play an active role in reunions with their mothers,'' Anim. Behav. **41**, 364–366.

Heffner, R. S., and Heffner, H. E. (**1986**). ''Localization of tones by horses: Use of binaural cues and the role of the superior olivary complex.'' J. Acoust. Soc. Am. **100**, 93–103.

Heffner, R. S., and Heffner, H. E. (**1990**). ''Vestigial hearing in a fossorial mammal, the pocket gopher (*Geomys bursarius*),'' Hear. Res. **46**, 239–252.

Heffner, R. S., and Heffner, H. E. (**1992a**). ''Evolution of sound localization in mammals,'' in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 691–715.

Heffner, R. S., and Heffner, H. E. (**1992b**). ''Hearing in large mammals: sound-localization acuity in cattle (*Bos taurus*) and goats (*Capra hircus*),'' J. Comp. Psychol. **106**, 107–113.

Heffner, R. S., Koay, G., and Heffner, H. E. (**2001**). ''Sound localization in a new-world frugivorous bat, *Artibeus jamaicensis*: Acuity, use of binaural cues, and relationship to vision,'' J. Acoust. Soc. Am. **109**, 412–421.

Holt, M. M., and Schusterman, R. J. (**2002**). ''Seals, sequences, and signal detection,'' Marine Mammal Sci. **18**, 208–212.

Insley, S. J. (**2001**). ''Mother-offspring vocal recognition in northern fur seals is mutual but asymmetrical,'' Anim. Behav. **61**, 129–137.

Kastak, D., and Schusterman, R. J. (**1996**). ''Temporary threshold shift in a harbor seal (*Phoca vitulina*),'' J. Acoust. Soc. Am. **100**, 905–1908.

Kastak, D., and Schusterman, R. J. (**1998**). ''Low frequency amphibious hearing in pinnipeds: Methods, measurements, noise, and ecology,'' J. Acoust. Soc. Am. **103**, 2216–2228.

Kastak, D., Schusterman, R. J., Southall, B. L., and Reichmuth, C. J. (**1999**). ''Underwater temporary threshold shift induced by octave-band noise in three species of pinniped,'' J. Acoust. Soc. Am. **106**, 1142–1148.

Masterton, B., Thompson, G. C., Bechtold, J. K., and Robards, M. J. (**1975**). ''Neuroanatomical basis of binaural phase-difference analysis for sound localization: A comparative study,'' J. Comp. Physiol. Psychol. **89**, 379–386.

Macpherson E. A., and Middlebrooks, J. C (**2002**). ''Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited,'' J. Acoust. Soc. Am. **111**, 2219–2236.

Mills, A. W. (**1958**). ''On the minimum audible angle,'' J. Acoust. Soc. Am. **30**, 237–246.

Møhl, B. (**1964**). ''Preliminary studies and hearing in seals,'' Vidensk. Medd. Dansk. Naturh. Foren. **127**, 283–294.

Moore, P. W. B. (**1975**). ''Underwater localization of click and pulsed pure-tone signals by the California sea lion (*Zalophus californianus*),'' J. Acoust. Soc. Am. **57**, 406–410.

Moore, P. W. B., and Au, W. L. (**1975**). ''Underwater localization of pulsed pure tones by the California sea lion (*Zalophus californianus*),'' J. Acoust. Soc. Am. **58**, 721–727.

Perry, E. A., and Renouf, D. (**1988**). ''Further studies of the role of harbour seal (*Phoca vitulina*) pup vocalizations in preventing separation of mother-pup pairs,'' Can. J. Zool. **66**, 934–938.

Peterson, R. S., and Bartholomew, G. A. (**1969**). ''Airborne vocal communication in the California sea lion,'' Anim. Behav. **17**, 17–24.

Petrinovich, L. (**1974**). ''Individual recognition of pup vocalizations by northern elephant seal mothers,'' Z. Tierpsychol **34**, 308–312.

Ravizza, R. J., and Masterton, B. (**1972**). ''Contribution of neocortex to sound localization in opossum (*Didelphis virginiana*),'' J. Neurophys. **35**, 344–356.

Renaud, D. L., and Popper, A. N. (**1975**). ''Sound localization by the bottlenose porpoise, *Tursiops truncates*,'' J. Exp. Biol. **63**, 569–585.

Richardson, J. W., Greene, S. R., Malme, C. I., and Thomson, D. H. (**1995**). *Marine Mammals and Noise* (Academic, New York).

Riedman, M. (**1990**). *The Pinnipeds: Seals, sea lions, and walruses* (Univ. of California, Berkeley, CA).

Schusterman, R. J. (**1978**). ''Vocal communication in pinnipeds,'' in *Behavior of Captive Wild Animals*, edited by H. Markowitz and V. J. Stevens (Nelson-Hall, Chicago, IL), pp. 247–308.

Schusterman, R. J., Kastak, D., Levenson, D. H., Reichmuth, C. J., and Southall, B. L. (**2000**). ''Why pinnipeds don't echolocate,'' J. Acoust. Soc. Am. **107**, 2256–2264.

Southall, B. L. (**2000**). ''Masking in three pinnipeds: Underwater, low-frequency critical ratios,'' J. Acoust. Soc. Am. **108**, 1322–1326.

Southall, B. L. (**2002**). ''Northern elephant seal field bioacoustics and aerial auditory masked hearing in three pinnipeds,'' unpublished doctoral dissertation, University of California, Santa Cruz, CA.

Southall, B. L., Schusterman, R. J., and Kastak, D. (**2003**). ''Auditory masking in three pinnipeds: aerial critical ratios and direct critical bandwidth measurements,'' J. Acoust. Soc. Am. **114**, 1660–1666.

Stirling, I., and Archibald, W. R. (**1977**). ''Aspects of predation of seals by polar bears,'' J. Fish. Res. Board Can. **34**, 1126–1129.

Strutt, J. W. (**1907**). ''On our perception of sound direction,'' Philos. Mag. **13**, 214–232.

Terhune, J. M. (**1974**). ''Directional hearing of a harbor seal in air and water,'' J. Acoust. Soc. Am. **56**, 1862–1865.

Trillmich, F. (**1981**). ''Mutual mother-pup recognition in Galapagos fur seals and sea lions: Cues used and functional significance,'' Behaviour **78**, 21–42.

Turnbull, S. D., and Terhune, J. M. (**1990**). ''White noise and pure tone masking of pure tone thresholds of a harbor seal listening in air and under water,'' Can. J. Zool. **68**, 2090–2097.

Wightman, F. L., and Kistler, D. J. (**1992**). ''The dominant role of low frequency interaural time differences in sound localization,'' J. Acoust. Soc. Am. **85**, 868–1661.

Yost, W. A. (**2000**). *Fundamentals of Hearing: An Introduction*, 4th ed. (Academic, San Diego, CA).

Tolerance by ringed seals (*Phoca hispida*) to impact pipe-driving and construction sounds at an oil production island[a)]

Susanna B. Blackwell[b)]
Greeneridge Sciences Inc., 1411 Firestone Road, Goleta, California 93117

John W. Lawson[c)]
LGL Limited, Environmental Research Associates, 22 Fisher Street, P.O. Box 280, King City, Ontario, L7B 1A6, Canada

Michael T. Williams
LGL Alaska Research Associates Inc., 1101 E. 76th Avenue, Suite B, Anchorage, Alaska 99518

(Received 1 July 2003; revised 15 November 2003; accepted 31 January 2004)

During June and July 2000, impact pipe-driving sounds at Northstar Island (Prudhoe Bay, Alaska) were recorded underwater and in air at distances 63–3000 m from the source. Simultaneously, reactions of nearby ringed seals (in water or on ice) were documented. Pipe-driving pulses were analyzed to determine unweighted peak and rms sound-pressure levels (SPLs) and sound-exposure levels (SELs). Underwater, mean levels for these parameters reached 157 and 151 dB *re*: 1 μPa and 145 dB *re*: 1 μPa^2·s, respectively, at 63 m. The corresponding values in air were 112 and 96 dB *re*: 20 μPa and 90 dB *re*: (20 μPa)^2·s, respectively. Underwater SPLs were $<$180 dB *re*: 1 μPa at all distances. During 55 h of observation, 23 observed seals exhibited little or no reaction to any industrial noise except approaching Bell 212 helicopters. Ringed seals swam in open water near the island throughout construction activities and as close as 46 m from the pipe-driving operation. Based on current audiometric data for seals, these sounds are expected to be audible to less than 3 km underwater and at least 0.5 km in air. Most likely the seals around Northstar Island were habituated to industrial sounds. © *2004 Acoustical Society of America.*
[DOI: 10.1121/1.1701899]

PACS numbers: 43.80.Nd, 43.50.Rq [WA] Pages: 2346–2357

I. INTRODUCTION

During the winter of 1999–2000, BP Exploration (Alaska) Inc. began construction of the Northstar oil production island ~5 km offshore of Long Island, northwest of Prudhoe Bay, Alaska, in the western Beaufort Sea. As part of the construction effort, an impact pipe-driver was employed to drive 35 well pipes into the gravel within the artificial island. Impact pipe-driving in or near the sea is known to produce strong pulses of underwater sound. For example, Greene (1999) reports on sounds from a Menck MHU 3000 hydraulic hammer driving platform pilings on the Scotian Shelf. Median broadband (20–5000 Hz) peak and rms sound-pressure levels, and sound-exposure levels, for pulses ($n=631$) received 1.5 km from the piles were 177 dB *re*: 1 μPa, 165 dB *re*: 1 μPa, and 156 dB *re*: 1 μPa^2·s, respectively. (All values quoted in this paper are unweighted.) The corresponding 95th percentile values were 179, 172, and 162 dB, respectively. In another study (Malme *et al.*, 1998), the calculated broadband (9–4490 Hz) source energy level for a pipe-driver operating in coastal shelf conditions was estimated to be 206 dB *re*: 1 μPa^2·s-m (Malme, personal communication, 2002). Würsig *et al.* (2000) report mean broadband (100–25 600 Hz) rms levels of 170 dB *re*: 1 μPa at 250 m, for 20 2.5-s samples each containing 1–2 pipe-driving blows. Finally, Johnson *et al.* (1986, also cited in Richardson *et al.*, 1995) measured underwater sounds with a cabled hydrophone placed 1000 m from Sandpiper Island (located ~25 km northwest of Northstar), where a 76-cm conductor pipe was being driven. The peak pressure level measured ca. 1 km away was 138 dB *re*: 1 μPa when the conductor pipe depth was 21 m. The strongest components were at 30–40 Hz and ca. 100 Hz.

Given these previous studies, it was predicted that sounds produced by impact driving of the pipes would be the strongest sounds during Northstar construction. There was concern that underwater sound pulses might be of sufficient intensity to cause temporary hearing impairment (i.e., Kastak *et al.* 1999) or other physical effects on nearby marine mammals, in particular the ringed seals (*Phoca hispida*) that are fairly common around the island (Kelly and Quakenbush, 1990; Moulton *et al.* 2002). The U.S. National Marine Fisheries Service (NMFS) has specified that pinnipeds and cetaceans should not be exposed to pulsed sounds at received levels exceeding, respectively, 190 and 180 dB *re*: 1 μPa on an rms basis, i.e., averaged over the pulse duration (NMFS, 2000).

Existing data on the behavioral reactions of pinnipeds to pulsed sounds suggest that ringed seals would exhibit either

[a)]Portions of this work were presented in "Tolerance of ringed seals (*Phoca hispida*) to sounds from impact pile-driving at an oil production island," ECOUS (Environmental Consequences Of Underwater Sound) meeting, San Antonio, TX, 12–16 May 2003.
[b)]Electronic mail: susanna@greeneridge.com
[c)]Current affiliation: Marine Mammals Section, Department of Fisheries and Oceans, P.O. Box 5667, Northwest Atlantic Fisheries Center, St. John's, Newfoundland, A1C 5X1 Canada.

no, or only short-term, behavioral responses even when exposed to strong pulsed sounds. Ringed seals in the Prudhoe Bay region often tolerated exposure to high received levels (180–190 dB re: 1 μPa rms) of low-frequency sound pulses from airgun arrays, with little evidence of changes in behavior and no more than localized avoidance (Harris *et al.*, 2001; Moulton and Lawson, 2002). However, it should be noted that sound pulses from pipe-driving are closer together than those from marine seismic surveys (i.e., interpulse intervals are typically 1.2–1.7 s vs 10–18 s).

The goals of this study were: (1) to characterize the acoustic properties of pipe-driving sounds underwater and in air at various distances from Northstar Island; and (2) to monitor the potential disturbance of pipe-driving sounds to ringed seals by recording the seals' reactions to the sounds. BP was prepared to suspend pipe-driving operations if underwater sound levels near seals were $\geqslant$190 dB re: 1 μPa (rms), the "safety level" specified by NMFS (2000). Seals within view of the island were therefore observed during pipe-driving activities until underwater recordings determined that the 190-dB level was not exceeded.

II. METHODS

A. Acoustical measurements and analyses

1. Field recordings of sounds in air and underwater

In June and July 2000, a diesel-powered Delmag D62-22 impact driver (36–50 blows/min, hammer weight=12 930 kg, energy per blow=224 kJ) was used to drive a total of 35 pipes below the surface of Northstar Island [Figs. 1(a), (b)]. The driver was lifted by detonation of diesel fuel in a cylinder, and then dropped ~3.4 m onto the pipe. The pipes included 25 well conductors and five disposal well pipes that were 51 cm in diameter, and five insulator pipes that were 107 cm in diameter. All pipes were driven individually, and disposal pipes were positioned concentrically to insulator pipes. In four out of five cases the disposal pipes were driven first. Simultaneously, the island side slopes were graded and concrete-block slope protection was placed from the edge of the sheet pile wall to a water depth of ~5 m [Fig. 1(b)]. The slope in this area was 1:3 down to the subsea berm [Fig. 1(b)] and then 1:2 down to the seafloor. Underwater and in-air recordings of impact pipe-driving sounds were obtained on 21 and 23 June for 51-cm pipes, and on 20 July for 107-cm pipes.

In June, when sounds from driving 51-cm pipes were recorded, ice breakup was just beginning and a moat of open water up to 10 m wide surrounded the island. The inner edge of the moat consisted of the island's concrete mats. The seaward edge of the moat consisted of continuous landfast sea ice that was 1.5–2.4 m thick. The maximum water depth in the moat was 6 m as the moat was above the sloping base of the island and the associated subsea berm [Fig. 1(b)]. The 51-cm pipes were driven at a location 43 m from the water's edge on the east side of the island [Fig. 1(a)]. A calibrated hydrophone and microphone were deployed successively at six stations on the ice. The closest recording station was 63 m from the pipe-driving, at the outer edge of the moat. At the further sites, 210, 390, 650, 1400, and 2960 m from the

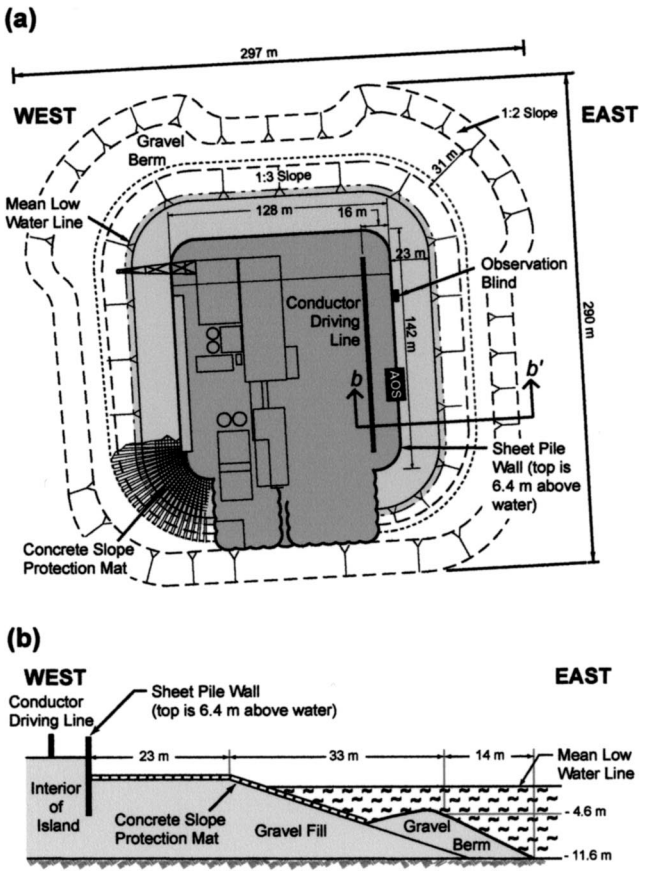


FIG. 1. (a) Schematic overview of Northstar Island showing the line within the eastern interior where pipes were driven (the island is slightly tilted on its E-W axis). Most observations were conducted from an alternate observation site (AOS) atop the eastern outer wall. On 28–29 June, observations were conducted from an observation blind positioned at the eastern outer wall. (b) Cross section of the eastern island edge and the underwater gravel berm offshore of the concrete mat slope protection, shown by the line $\overline{bb'}$ in (a).

driven pipes [Fig. 2], the hydrophone was deployed through an augered hole in the ice or through a seal breathing hole. The recording locations were to the east or east–southeast, as the strongest waterborne sounds were expected in that direction. Water depths at the recording stations were 11–12 m (except in the moat), and the hydrophone depth was always 10 m (except in the moat where it rested on the bottom). Recording locations were determined with a laser rangefinder (Bushnell Yardage Pro 20-0880) at short distances (<800 m) and a GPS receiver (Garmin 12XL) at longer distances from the island.

On 20 July recordings were made during impact driving of 107-cm well insulator pipes. The five recording stations, 80, 160, 200, 500, and 1110 m from the insulator pipes [Fig. 2], were reached by boat, as the sea ice had receded offshore (the short duration of the pipe-driving session on that day did not give us time to make measurements at greater distances). Once the vessel was in position at a recording station, the hydrophone was suspended directly from the vessel's side into the water at a depth of 9 m. A depth reading was taken with the vessel's depth sounder, and then all sound sources (engines, generator, depth sounder) on the vessel were turned off. The microphone was positioned on the deck of the vessel

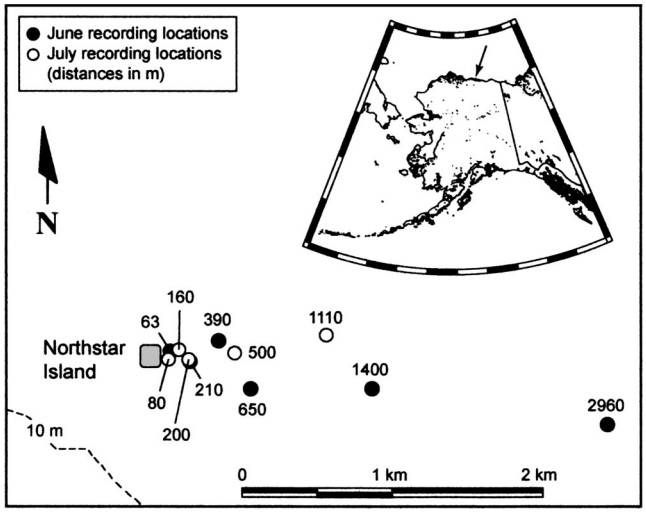


FIG. 2. Acoustic recording locations during impact hammering of 51-cm well conductor pipes in June (filled circles) and 107-cm well insulator pipes in July 2000 (open circles), at Northstar Island. The island is indicated by a gray polygon. Recording locations are labeled with their respective distances (in m) to the pipe being driven.

in such a way that it had an unobstructed path to the pipes at all times. Recording locations were determined as in June.

The impact pipe-driver could be operated at two energy levels. We report the acoustic data for the driver operated at ''level 4,'' the highest energy level available from this impact driver.

In June, background sound measurements were made between pipe-driving sessions at the farthest distance (2960 m) from the island, in the absence of other obvious strong sound sources. The results do not represent true ''ambient'' measurements, as distant construction sounds from Northstar were included on the recordings.

2. Acoustic equipment

The hydrophone was an International Transducer Corporation (ITC) model 6050C, which includes a low-noise preamplifier next to the sensor and a 30-m cable. Prior to recording, the hydrophone signals were amplified with an adjustable-gain postamplifier. The omnidirectional microphone was an ACO model 7013 condenser microphone with a 4012 preamplifier and a windscreen. Signals from both sensors were recorded simultaneously on two channels of a Sony model PC208Ax instrumentation-quality digital audio tape (DAT) recorder, at a sampling rate of 24 kHz. Quantization was 16 bits, providing a dynamic range of >80 dB between an overloaded signal and the quantization noise. Date and time were recorded automatically. The field acoustician noted the industry activities during each recording on a memo channel.

3. Signal analysis

The recorded, digitized hydrophone and microphone signals were transferred directly to a computer hard drive as time series. They were then equalized and calibrated in units of sound pressure with flat frequency response over the data bandwidth (4–10 000 Hz). Hydrophone and microphone signals were unweighted and analyzed in the same way, except

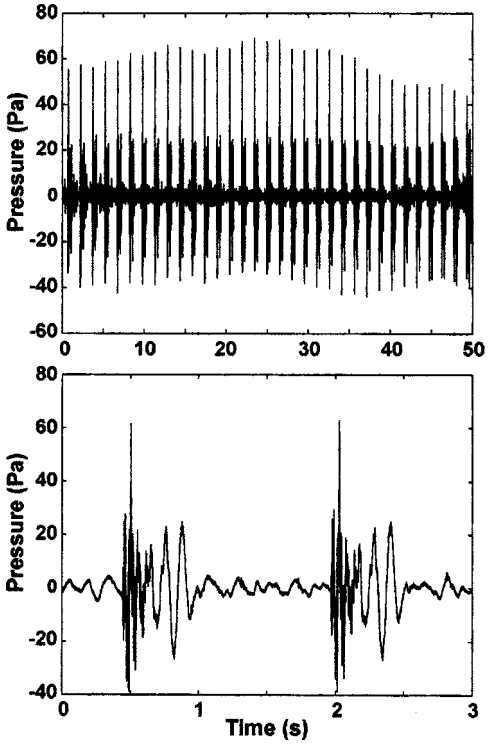


FIG. 3. Sound-pressure time series from an underwater recording of pulses during driving of 107-cm pipe, July 2000. (a) Fifty seconds of recording at a mean distance of 80 m from the pipe. (b) Detail of two pulses (over 3 s) from the same distance.

the microphone data are referenced to 20 μPa. Analysis was done using MATLAB (The Math Works, Natick, MA) routines and custom programs. For each recording, a sound-pressure time series (waveform) was generated to help select samples for further analysis; two examples are shown in Fig. 3. Impact driving pulses were analyzed using time series analysis (TSA) routines previously developed for transient pulses generally <1 s long, such as airgun pulses in seismic surveys (Burgess and Greene, 1999; McCauley *et al.*, 1998, 2000, also used similar measures). TSA procedures determined four parameters associated with each pulse.

(1) *peak pressure*, i.e., the instantaneous maximum of the absolute sound pressure, in dB re: 1 μPa (underwater) or dB re: 20 μPa (in air);
(2) pulse *duration*, defined as the time interval between the arrival of 5% and 95% of the total pulse energy, in s;
(3) pulse *sound-pressure level* (SPL), averaged over the pulse duration, in dB re: 1 μPa or dB re: 20 μPa; and
(4) pulse *sound-exposure level* (SEL), defined as the squared instantaneous sound pressure integrated over the pulse duration, in dB re: 1 $\mu\text{Pa}^2\cdot$s or dB re: $(20\ \mu\text{Pa})^2\cdot$s. This measures the energy in the pulse, excluding the background sound energy as characterized by measurements between pulses.

SPL values can be influenced significantly by the duration, depending on the pulse shape. For example, low-level reverberation that extends the pulse length can result in a longer averaging time and thus a lower SPL. The SEL is not influ-

TABLE I. Summary of observation effort, number of ringed seals sighted, and construction activities at Northstar Island during the visual monitoring program in June and July 2000.

Date	Observation duration (hr)	Number of ringed seals observed	Construction and other activities
19 June	7	One on ice at 1 km	Slope preparation; no pipe-driving, helicopter traffic
20 June	4 (from helicopter and on ice)	Adult on ice; one of unknown age swimming	Vibrating 51-cm pipes; slope preparation; drilling sound recording holes in ice, helicopter traffic
21 June	10.5	Juvenile swimming; adult on ice at 500 m; adult on ice at 1.5 km; one unknown age swimming; two adults on ice at 1–2 km	Two 51-cm pipes driven; slope protection; helicopter traffic
22 June	3.5	None sighted	Slope protection; no pipe-driving, helicopter traffic
23 June	0	One of unknown age swimming	51-cm pipes driven; slope protection; helicopter traffic
26 June	0	No observation data	Preliminary 107-cm pipe-driving; slope protection; helicopter traffic
27 June	8.25	Two swimming; one on ice 1 km to E	107-cm pipes driven; slope protection, helicopter traffic
28 June	5	Three adults hauled on ice at least 1 km to E	107-cm pipes driven; slope protection, helicopter traffic
29 June	8.5	Seven on ice at least 1 km to E and SE	107-cm pipes driven; slope protection, helicopter traffic
20 July	8	None sighted	107-cm pipes driven; slope protection; crew vessel traffic
Total	54.75	23	

enced by the duration, as it is a measure of the total energy in the pulse regardless of the pulse duration.

The frequency composition of the pulses was determined using power spectrum density routines. The hydrophone data were high-pass filtered at 10 Hz to remove background noise that hampered analysis of the pulses received at the more distant stations. However, to characterize the contribution of frequencies below 10 Hz to underwater sounds, we report the difference between raw and high-pass filtered data for one of the closer stations (160 m). Microphone data were not high-pass filtered.

Broadband background sound levels were obtained by analyzing samples 8.5 s long from times with no pipe-driving (on the same days). Frequency composition was determined by calculating the sound-pressure spectral density by Fourier analysis, using the Blackman–Harris minimum three-term window (Harris, 1978). The transform length was 1 s. With windowing, the spectral resolution was 1.7 Hz with 1-Hz bin separation. Transforms were overlapped by 50% and the resultant 16 power spectral densities were averaged for each 8.5-s sample.

A simple propagation model was fitted by the least-squares method to the peak, rms, and SEL values from both the hydrophone and microphone, in order to develop equations that characterize propagation loss underwater and in air. We used two different models, based on logarithmic spreading loss

$$\text{Received level } (\mathbf{RL}) = A - B \cdot \log(R), \tag{1}$$

and

$$\mathbf{RL} = A - B \cdot \log(R) - C \cdot R. \tag{2}$$

In these equations, R is the range in m and the units for **RL** are dB re: 1 μPa for underwater SPLs and dB re: 20 μPa for in-air SPLs. The constant term (A) is the hypothetical extrapolated level at distance 1 m based on far-field measurements. This hypothetical value would equal the actual level at 1 m only if the source were a point source and if loss rates were consistent at all distances from 1 m to the maximum measurement distance, neither of which is the case, at least for the underwater sounds. The estimated A value is useful mainly as a basis for comparison with other sound sources operating in the same region. The spreading loss term (B) varies with the dominant frequency in the pulse, water depth, bottom topography and composition, and presence of ice cover. C, in dB/m, is the coefficient describing losses due to scattering and absorption. Roughness of the underside of the ice or the composition and topography of the seabed will affect such losses. Equation (1) is a simplification of Eq. (2) and is useful when there are few measured distances or when the greatest measurement range is short, less than 1 km.

B. Seal behavioral observations

Behavioral observations at Northstar provided information on seal distribution, abundance, and behavior during periods with and without impact pipe-driving and other construction activities. Observations were done on 19–23 June, 26–29 June, and 20 July 2000, from one of two sites: either an elevated position on the sheet pile wall [AOS in Fig. 1(a), eye height=8.2 m above sea level], or an enclosed observation hide hung on the outside of the sheet pile wall [see Fig. 1(a), observation height=3.8 m above sea level]. Weather conditions were generally ideal for marine mammal observations; light winds and warm temperatures allowed observers

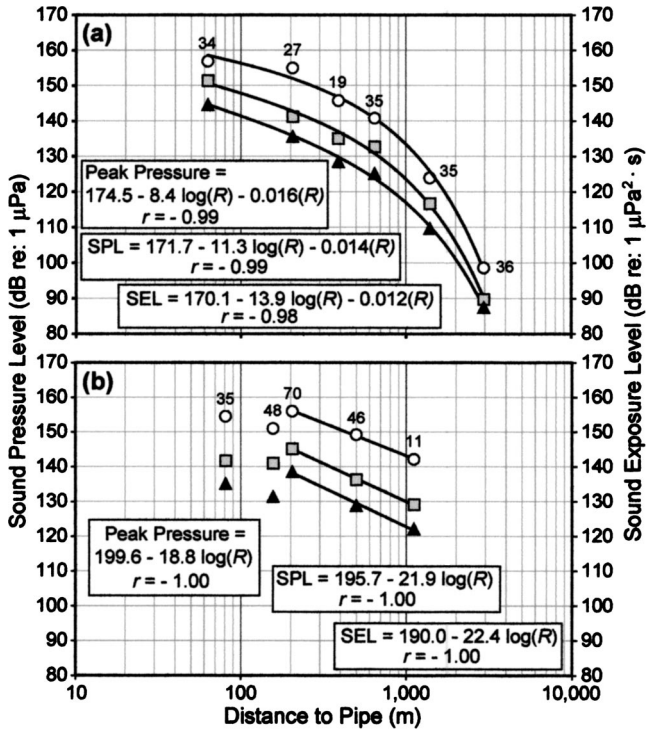


FIG. 4. Mean levels of underwater sound pulses received during impact driving of (a) 51-cm well conductor pipes in June and (b) 107-cm well insulator pipes in July 2000. Shown are peak pressure (white circles) and rms SPL (gray squares), with units to be read on left y axis, in dB re: 1 μPa; and pulse SEL (black triangles), with units to be read on right y axis, in dB re: 1 μPa^2·s; bandwidth 10–10 000 Hz, unweighted. Sample size is shown above top symbol for each recording station. Standard deviation bars are not shown for each mean for the sake of clarity; in June s.d. was always <1.0 dB and on average 0.5 dB; in July s.d. was always <3.2 dB and on average 1.1 dB.

to remain at the elevated observation site on all days except 28–29 June. Daily observation periods lasted from 3.5 to 10.5 h, for a total of 54.75 h during the 10 days [Table I]; 14.5 h occurred on days when there was no impact driving. On days with pipe-driving, daily observations began shortly before the impact driving commenced, and continued after driving operations ceased for the day, for a total of 40.25 h. Days with pipe-driving therefore included periods with active driving as well as periods during which the pipes were positioned, the pile-driver was moved, etc. Driving was never continuous during the entire observation period. Observer safety and logistical restrictions imposed limits on the authors' time on the island, and hence the amount of observation data collected during periods when pipes were not being driven. The observation effort was therefore primarily a means to mitigate potential impacts on ringed seals, rather than an experiment to compare seals' behavior during periods with and without impact pipe-driving.

In June 2000, primary construction activities on Northstar included preparation of the island's side slopes outside the sheet pile walls, lifting and laying the concrete mat slope protection, emplacement of module footings, pipe-driving, and regular helicopter traffic to and from the island. Marine mammal observations were conducted during the above construction activities, and during quiet periods before and/or after these activities occurred.

The 51-cm pipes were driven intermittently from 21 to 23 June, and again in late July. The 107-cm pipes were driven intermittently from 26 to 29 June and again starting on 20 July. Ringed seals were observed during each day with impact pipe-driving until the acoustical measurements had confirmed that the 180 and 190 dB (re: 1 μPa, rms) safety radii for both pipe diameters were *within* the island, i.e., underwater SPLs were below these values at all distances.

Observers determined the locations of seals relative to the observation site and any potential source of disturbance by knowing the distances to nearby visual landmarks on the ice edge surrounding the open-water moat, as measured using a Bushnell Yardage Pro 800 laser rangefinder. During periods of *ad lib* observation, the marine mammal monitors recorded the following:

(1) beginning and end times of construction activities;
(2) number of seals seen;
(3) bearing and distance between the seal(s) and the activities;
(4) behavior of the seal(s), in particular any indications of disturbance or reactions to construction or other activities.

In describing seal behavior, the observers made a subjective assessment of the rate (''pace'') at which the seal responded to an anthropogenic activity. Seals that appeared to be relatively relaxed or were moving slowly were coded as ''sedate.'' Seals that appeared to be agitated, moved frantically, or dived or moved rapidly away from the source of activity were classified as behaving ''vigorously.'' Intermediate states were recorded as ''moderate.''

III. RESULTS

A. Sound measurements

1. Underwater sounds

Figure 4 shows measured broadband levels of underwater sound during impact driving of (a) 51-cm pipes in June and (b) 107-cm pipes in July. The sound propagation model represented by Eq. (2) was fitted to the June data [Fig. 4(a)]. Similarly, Eq. (1) was used for the July data [Fig. 4(b)]. Measurements obtained less than 200 m from the island in July were not used in fitting the model. They were unexpectedly low, indicative of either lower source levels or different propagation processes than at distances $\geqslant$200 m. Consequently, in July the regression included only three sampling stations, covering a range of distances equal to only 27% of the range used for fitting the model in June. The highest values for each parameter were obtained at the closest station (63 m) in June, but at the 200-m station in July. The overall highest recorded levels were obtained 63 m from a 51-cm pipe in June: peak level averaged 157 dB re: 1 μPa (maximum 158 dB), rms SPL averaged 151 dB re: 1 μPa (max. 152 dB), and SEL averaged 145 dB re: 1 μPa^2·s (max. 146 dB). In June, when most of the propagation path was under shorefast ice, all sound parameters decreased rapidly with distance. In addition to spreading losses [B in Eq. (1)] of 8.4–13.9 dB/tenfold change in distance, scattering losses (C) resulted in marked losses at longer distances from the island. For example, with $C=0.014$ dB/m (or 14 dB/km), the re-

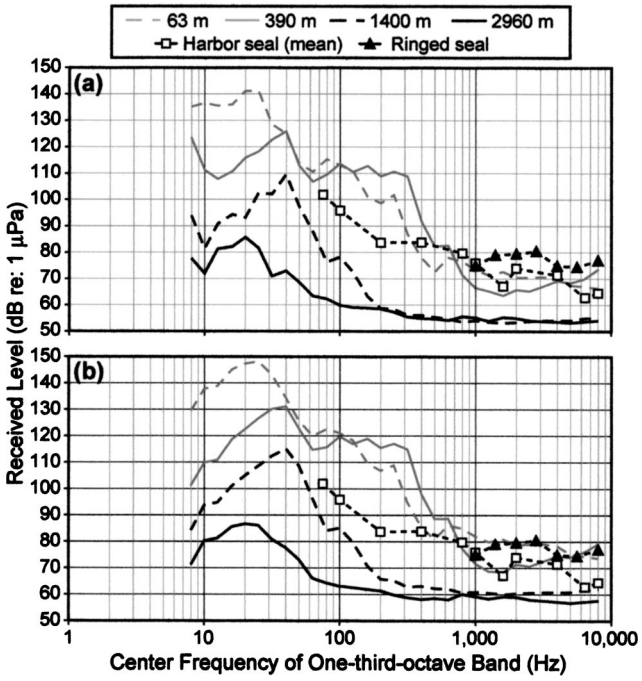


FIG. 5. One-third-octave band levels for underwater pulses recorded in June, calculated (a) by spectral analysis of an 8.5-s sample comprising several pipe-driving pulses; and (b) by calculating SPL from SEL and pulse duration for each pulse within the above 8.5-s sample and plotting the mean value. Values from four stations (63, 390, 1400, and 2960 m) are plotted for each method. Also shown are an underwater audiogram for the ringed seal (Terhune and Ronald, 1975; black triangles) and a composite audiogram for the harbor seal from three independent measurements (Møhl, 1968; Terhune and Turnbull, 1995; Kastak and Schusterman, 1998; white squares).

ceived SPL at distance 3 km was 90 dB *re*: 1 μPa. In July, when there was little ice, spreading loss terms (over a narrow span of distances) were close to what one would expect for spherical spreading, 19–22-dB/tenfold change in distance. At the farthest July station (1110 m), rms SPL was ~130 dB *re*: 1 μPa; if the simpler Eq. (1) was applied to the June data (over the range of distances 210–1400 m), the rms SPL at 1110 m would be 128 dB *re*: 1 μPa, i.e., almost the same value as in July. Unfortunately, measurements beyond 1110 m could not be obtained in July.

Mean pulse durations were 0.18–0.60 s in June and 0.11–0.22 s in July. Pulse durations were positively correlated with distance from pipe-driving both in June ($r=0.80$, $df=184$, $p<0.01$) and July ($r=0.29$, $df=84$, $p<0.01$). The smaller r value in July is likely the result of a much smaller range of distances, ~420 m vs ~2900 m in June. The longest pulse durations were obtained at the farthest stations in both months.

The 63-m station in June, at the outer edge of the moat, was only ~10 m from Northstar's shore. Therefore, calculated distances for underwater SPLs of 190 and 180 dB, above which impulsive sounds are considered to be a concern for marine mammal safety (NMFS, 2000), were *within* the confines of the island.

Figure 5 shows one-third-octave band SPLs for pulses recorded in June at four stations: 63, 390, 1400, and 2960 m from the driven pipe. Band levels were calculated two different ways: (a) by considering the pipe-driving sound to be a ''constant'' sound source and doing spectral analysis on 8.5-s samples (one per station), each comprising 6–7 pulses [Fig. 5(a)]; and (b) by calculating an SPL from the one-third-octave band SEL of each individual pulse included in the 8.5-s samples (taking into account the pulse duration), and then plotting the average value for each one-third-octave band [Fig. 5(b)]. With the latter technique, the values correspond to what one would obtain with a ''continuous'' sound consisting only of the pulses themselves, without the interpulse intervals. This approimates the upper limit of SPLs that a seal might perceive. Note that the one-third-octave levels are highest at the lowest frequencies analyzed, and diminish with increasing frequency up to 100–1000 Hz, depending on range. The levels at all ranges flatten for frequencies >100–1000 Hz, suggesting that the measured levels at these frequencies are background noise at the recording site (which includes island noise), not pipe-driving impact noise. Figure 5 also shows in-water auditory thresholds for ringed seals and harbor seals (*Phoca vitulina*); see Sec. IV for a comparison of the measured one-third-octave levels with the auditory thresholds.

Pipe-driving pulses included sound components below 10 Hz. To assess the effect of high-pass filtering on the results, 24 pulses recorded at the 160-m station on 20 July were analyzed both with and without 10-Hz high-pass filtering. Without filtering, peak levels increased by 0.35 ± 0.04 dB (mean ±1 s.d.), SEL increased by 1.81 ± 0.1 dB, pulse duration increased by 0.52 ± 0.02 s, and (consequently) the rms SPL measured over the pulse duration decreased by 5.52 ± 0.21 dB. Generally, over 95% of the energy in pipe-driving pulses was below 224 Hz (the upper limit of the one-third-octave band centered at 200 Hz), both in June and July. In the data collected at 63 m in June, nearly 99% of the energy was below 71 Hz (one-third-octave band centered at 63 Hz), but this was the only station at which the hydrophone was lying on the bottom instead of being suspended in the water column.

The lower limit of broadband (10–10 000 Hz) background levels were ~85 and ~125 dB *re*: 1 μPa in June (at 3 km) and July (at 1.1 km), respectively. In July, island sounds were dominant on the recording, even in the absence of pipe-driving. The lower June background levels were probably attributable to the lack of ice in July. In addition to altering the propagation conditions, the July open-water conditions led to vessel traffic (crew boat and barges).

2. In-air sounds

Figure 6 shows measured broadband levels of in-air sound during impact driving of 51-cm pipes in June (small symbols) and 107-cm pipes in July (large symbols). Because of poor signal-to-noise ratios, only six stations yielded valid airborne measurements: 63 and 210 m in June, and 80, 160, 200, and 500 m in July. The sound propagation model represented by Eq. (1) was fitted to the July data obtained at distances >100 m; data from the closer (80 m) station were excluded because they were unexpectedly low. The highest levels were obtained at the closest (63-m) station in June and the 160-m station in July. At the latter, peak level averaged 111 dB *re*: 20 μPa (maximum 115 dB), rms SPL averaged 93 dB *re*: 20 μPa (max. 97 dB), and SEL averaged 87 dB

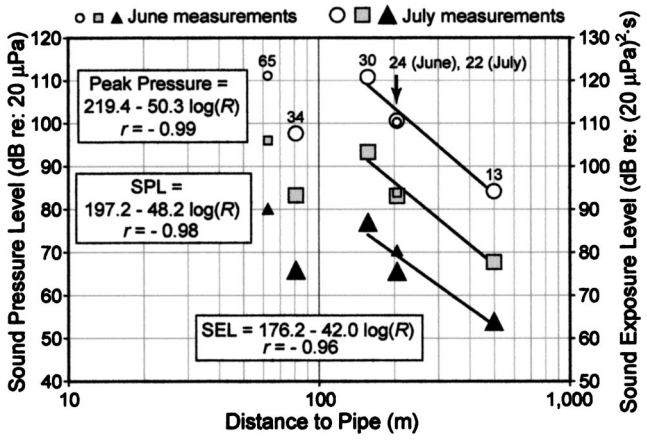


FIG. 6. Mean levels of in-air sound pulses received during impact driving of 51-cm well conductor pipes in June (small symbols) and 107-cm well insulator pipes in July 2000 (large symbols). Shown are peak pressure (white circles) and rms SPL (gray squares), with units to be read on left *y* axis, in dB *re*: 20 μPa; and pulse SEL (black triangles), with units to be read on right *y* axis, in dB *re*: (20 μPa)^2·s; bandwidth 10–10 000 Hz, unweighted. Sample size is shown above top symbol for each recording station. Standard deviation bars are not shown for each mean for the sake of clarity; in June s.d. was always <4.5 dB and on average 3.3 dB; in July s.d. was always <2.9 dB and on average 1.6 dB.

re: (20 μPa)^2·s (max. 90 dB). Spreading loss terms [*B* in Eq. (1)] were high in July, 42–50-dB/tenfold change in distance.

Mean pulse durations were 0.27–0.46 s in June and 0.17–0.63 s in July. Pulse durations were positively correlated with distance from the pipe-driving both in June ($r=0.50$, $df=105$, $p<0.01$) and July ($r=0.62$, $df=95$, $p<0.01$).

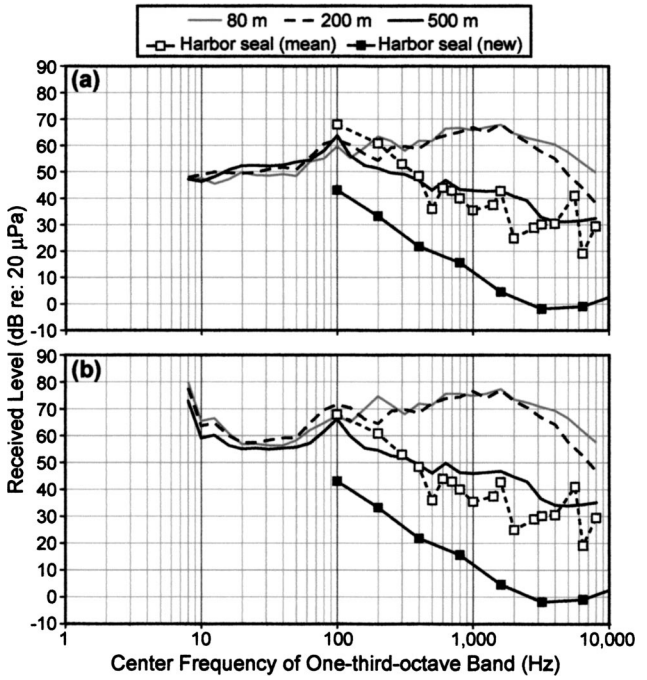


FIG. 7. One-third-octave band levels for in-air pulses recorded in July, calculated (a) by spectral analysis of an 8.5-s sample comprising several pipe-driving pulses; and (b) by calculating SPL from SEL and pulse duration for each pulse within the above 8.5-s sample and plotting the mean value. Values from three stations (81, 205, and 500 m) are plotted for each method. Two in-air audiograms for harbor seals are shown: a composite based on three independent measurements (Møhl, 1968; Terhune and Turnbull, 1995; Kastak and Schusterman, 1998; white squares) and more recent data collected by Schusterman *et al.* (2003) (black squares; see the text).

Figure 7 shows one-third-octave band SPLs for pulses recorded in July at three stations: 80, 200, and 500 m from the driven pipe. Band levels were calculated the same way as described for Fig. 5. Figure 7 also shows two sets of in-air auditory thresholds for harbor seals (see the caption of Fig. 7 and Sec. IV).

In contrast to the underwater data, the energy of in-air pipe-driving sounds was spread over a broader range of frequencies. In June, >95% of the energy was found in the frequency band 89–3534 Hz (one-third-octave bands centered at 100–3150 Hz). In July, the equivalent band was 143–5610 Hz (one-third-octave bands centered at 160–5000 Hz). During both months, the four strongest one-third-octave band levels were centered at 200, 630, 1000, and 1600 Hz, with the last being the highest.

The lowest measured level of in-air broadband background noise was ~68 and 65 dB *re*: 20 μPa in June and July, respectively (10–10 000 Hz, unweighted). These measurements were made at the farthest stations each month, in the absence of pipe-driving activity.

B. Seal reactions to sounds

During nearly 55 h of behavioral observations, 23 ringed seals were sighted around Northstar Island [Table II]. Most (17) were basking on the ice 0.5 to 2 km from the eastern edge of the island; the remaining seals (6) were seen swimming in the moat, 3–15 m from the outer edge of the island.

1. Pipe-driving sounds

Ringed seals near the island did not react strongly to pipe-driving activities. Twentey seals (15 basking on the ice and 5 in the moat) were observed during days with pipe-driving (totaling 40.25 h of observation).

(i) Eleven seals were seen without the concurrent presence of helicopter operations. Out of these, eight showed no apparent reaction to pipe-driving activities, one continued swimming, and two seals looked briefly at the island or pipe-driving operation before resuming their former activity.

(ii) Nine seals were seen at the same time as ongoing helicopter operations. None of these seals reacted to the pipe-driving, but they all reacted to some degree to the helicopter (see below and Table II).

Seals exhibiting no negative reactions to the pipe-driving included a juvenile that swam in the moat within 3 m of the water's edge along the island margin, directly offshore of the 51-cm pipe-driving operation, on 21 June. This seal appeared unafraid of the acoustic and visual stimuli associated with pipe-driving, and was observed making several approaches to investigate crews working near the water's edge.

2. Helicopter activities

Twelve seals were observed during the low overflight of a Bell 212 helicopter (during nine of those observations pipe-driving was taking place concurrently, see above). One seal showed no reaction to the helicopter. The remaining 11 reacted to the aircraft, either by increasing their vigilance and looking at the helicopter (10 seals) or by departing from a

TABLE II. Details of 23 ringed seal sightings that occurred during visual monitoring at Northstar Island during June and July 2000. Unk=unknown.

			Construction operations occurring			Seal behavior	
Date	Substrate	Distance from island	Pipe driving?	Helicopter ops?	Slope preparation?	Reaction	Pace
19 June	Ice	1 km	No	Yes	Yes	None apparent	Sedate
20 June	Ice	0.2 km	Vibratory	Yes	Yes	Looked at helicopter	Sedate
20 June	Moat	10 m	Vibratory	Yes	Yes	Looked at helicopter	Sedate
21 June	Moat	3 m	Yes: 51-cm	No	Yes	Looked at island	Sedate
21 June	Ice	0.5 km	Yes: 51-cm	Yes	Yes	Looked at helicopter	Moderate
21 June	Ice	1.5 km	Yes: 51-cm	No	Yes	None apparent	Sedate
21 June	Moat	10 m	Yes: 51-cm	No	Yes	None apparent	Sedate
21 June	Ice	1–2 km	Yes: 51-cm	No	Yes	None apparent	Sedate
21 June	Ice	1–2 km	Yes: 51-cm	Yes	Yes	Looked at helicopter	Sedate
23 June	Moat	15 m	Yes: 51-cm	Yes	Yes	Looked at helicopter	Sedate
27 June	Ice	1 km	Yes: 107-cm	No	Yes	Looked at onset of hammering	Sedate
27 June	Moat	Unk	Yes: 107-cm	No	Yes	Swimming	Sedate
27 June	Moat	Unk	Yes: 107-cm	Yes	Yes	Looked at helicopter	Sedate
28 June	Ice	1 km	Yes: 107-cm	No	Yes	None apparent	Sedate
28 June	Ice	1 km	Yes: 107-cm	No	Yes	None apparent	Sedate
28 June	Ice	1 km	Yes: 107-cm	Yes	Yes	Returned to water as helicopter landed	Moderate
29 June	Ice	2 km	Yes: 107-cm	Yes	Yes	Looked at helicopter	Moderate
29 June	Ice	1 km	Yes: 107-cm	Yes	Yes	Looked at helicopter	Sedate
29 June	Ice	2 km	Yes: 107-cm	Yes	Yes	Looked at helicopter	Sedate
29 June	Ice	1.5 km	Yes: 107-cm	Yes	Yes	Looked at helicopter	Sedate
29 June	Ice	1.5 km	Yes: 107-cm	No	Yes	None apparent	Sedate
29 June	Ice	1 km	Yes: 107-cm	No	Yes	None apparent	Sedate
29 June	Ice	1 km	Yes: 107-cm	No	Yes	None apparent	Sedate

basking site (1 seal). The disturbed seal was an adult seen basking on the fast ice 500 m from the SE corner of the island. This seal did not move during pipe-driving or in response to the proximity of the sound recording crew. However, it did return to the water when the helicopter circled over its haulout position, at a distance of ~100 m.

3. Other activities

At times when neither pipe-driving nor helicopter activities were occurring, seals did not appear to react to other industrial sound sources on the island. In almost all cases seals on the ice or in the water appeared to respond minimally or not at all to island-based construction activities. Basking seals glanced occasionally toward the island, and swimming seals watched people and equipment working on the island's margins. For example, construction crews reported seeing seals swimming nearby during slope preparation (these sightings are not included in Table II). During that activity, a large-bucket drag-line crane was used to deposit gravel onto the underwater berm [Fig. 1(b)] or the outer slope of the island margin prior to laying concrete mats.

IV. DISCUSSION

A. Acoustic characteristics of pipe-driving sounds

Levels reported in this study were lower than those reported previously for open-water pile-driving at comparable ranges (see the Introduction for a list of studies), most likely because of attenuation through the sediments deposited during island construction. There was also no clear relationship between pipe size and magnitude of the sound produced. At close range (63–81 m), underwater rms SPLs and SELs were higher for the 51-cm well conductor pipes. In air, all three parameters were higher for the smaller pipes. This difference was mainly due to the unexpectedly low values obtained at close range in July [Figs. 4 and 6]. At about 200 m, values for all three parameters were similar for both pipe sizes, both underwater and in air. Similarly, the highest peak values for the two pipe types, underwater and in air, were nearly the same (156–157 dB re: 1 μPa and 111–112 dB re: 20 μPa, respectively). This could suggest that the sound radiated was more a function of the material through which the pipes were being driven and the operating depth, rather than the size of the object being driven. However, we have no information on the substrate composition below the seafloor. In addition, the moat structure around the island [Fig. 1] could affect underwater measurements in the near field (<200 m), but should not affect in-air measurements. Consequently, the comparatively low values seen both in air and underwater during measurements close to the 107-cm pipe-driving [Figs. 4(b), 6] could have more to do with the source levels simply being lower at the time of sampling. Sound-pressure levels during pipe-driving can vary by as much as 18 dB solely on the basis of the substrate being penetrated and the increasing resistance over the course of a driving session (SBB, personal observation). Therefore, unless the recordings at all stations are made simultaneously, it is normal to see values from certain stations that do not conform to expectations. Simultaneous recordings at various distances were not practical in this study.

Underwater propagation loss apparently was higher in June than in July [Fig. 4]. This could be due to the reduced

effective water depth and possible increased transmission loss in June due to ice. However, direct comparisons are confounded by the difference in the span of recording distances during the two months.

Transmission loss (TL) varies with frequency and water depth; in shallow water, TL is high for both low and high frequencies (Richardson *et al.*, 1995). For the low frequencies in the underwater pulse spectrum and the shallow-water conditions around Northstar, cylindrical spreading $[10\cdot\log(R)]$ or mode stripping $[15\cdot\log(R)$, between cylindrical and spherical spreading] would be expected for all conditions. The June data conform to this prediction; in July, the smaller range of distances at which recordings were obtained (910 vs 2900 m) and the fewer data points (three vs six) for the logarithmic fit may have introduced some inaccuracy in fit within a simple TL model.

In-air propagation loss based on the July measurements was 42–50-dB/tenfold change in distance, whereas values of ~20-dB/tenfold change in distance (i.e., spherical spreading) would be expected for normal sound transmission in open air. Refraction caused by temperature or wind gradients could cause heightened spreading loss terms, but we have no information on the upper-air weather conditions during our measurements.

The frequency distribution of underwater and in-air pulses was quite different, but not unexpected [see Figs. 5 and 7]. Bjorno (1977) and Hamilton (1980) tested the relationship between sound attenuation (dB/m) and frequency (Hz) during propagation through components of the seafloor (i.e., mud, silt, clay, sand, etc.). They found an approximate first-power dependence of attenuation on frequency (i.e., a linear relationship): whereas the attenuation rate at 100 Hz is on the order of 0.1 dB/m, at 1 and 10 kHz it approaches 1 and 10 dB/m, respectively. Therefore, we can expect received levels for higher frequencies, generated when the metal hammer hits the metal pipe [see Fig. 7], to attenuate severely and be low even at 63 m, the closest station. The shallow-water depths (~11 m) at all recording stations should prevent sound transmission through the water column at frequencies below about 34 Hz, due to the modal cutoff when depth $\leqslant\frac{1}{4}$ wavelength (Officer, 1958). However, if the bottom material is porous the cutoff is not sharp because sound-pressure waves partially penetrate the bottom. This can be seem in the peak spectrum levels of 20–40 Hz shown in Fig. 5.

Greene and Burgess (2000) measured pulsed sounds from airgun arrays propagating in shallow water in the Prudhoe Bay area. Underwater pulse durations increased linearly with distance from the source, but only for distances >500 m at a hydrophone depth of 8 m. At shorter ranges their pulse durations were extremely variable. Only our June underwater dataset included measurements at distances >500 m. When pulse duration was regressed on distance at 650, 1400, and 2960 m the fit was tighter ($r=0.89$) than when using all stations ($r=0.80$; see Sec. III).

B. Sound levels received by ringed seals

The one-third-octave information provides a starting point for evaluating the frequency components of the pipe-driving sounds that ringed seals can hear, and the distances within which these components can be heard. Phocid audiogram data are scant below 100 Hz both underwater and in air, but harbor seals have deteriorating hearing sensitivity as frequency diminishes to 75–100 Hz (the lowest frequencies measured), and presumably have little or no sensitivity below 10 Hz. With >95% of the underwater energy below ~225 Hz, most of the pulse energy fell below the range of best sensitivity for *Phoca* underwater (~1–40 kHz, Richardson *et al.*, 1995). However, despite phocids' decreasing sensitivity at these frequencies, audiometric data [Fig. 5] suggest that underwater pulses (in June) were still audible at 390 m, and perhaps barely audible at 1400 m, but apparently inaudible at 2960 m. In air the pulse energy was spread over a wider range of frequencies and the one-third-octave spectrum peaked at a much higher frequency (near 1600 Hz vs ≤40 Hz underwater). More than 70% of the pulse energy was between 400 and 6300 Hz, which includes the range of best sensitivity for *Phoca* in air (~2–20 kHz, Richardson *et al.*, 1995). Available hearing data for *Phoca*, a ''composite'' of data collected by Møhl (1968), Terhune and Turnbull (1995), and Kastak and Schusterman (1998), would indicate that the in-air pulses measured in July were audible at 200 m but only weakly audible at 500 m. However, recent data collected by Schusterman *et al.* (2003) suggest these previous measurements of aerial hearing sensitivity were noise limited. The new data show acute aerial auditory thresholds that are close to human values in the range of best sensitivity (Sivian and White, 1933 via Fay, 1988). If in-air hearing of ringed seals is comparable, estimates of audible distances based on these data would indicate that the pulses were still clearly audible at 500 m. We do not have in-air recordings far enough from Northstar to determine the limit of pipe-driving noise audibility. However, at 1110 m the pulses were faintly audible to the field crew; considering the similarity between the new harbor seal audiogram and a human audiogram (see the reference above) it is likely that this distance is close to the limit of detection by the seals.

Another factor to consider is the effect of signal duration. Threshold levels have been shown to increase as signal duration decreases below a certain length (Johnson, 1968, 1991). In harbor seals, Terhune (1988, 1989) showed that thresholds to tone pulses were similar for durations above 50 ms, but increased as duration diminished below 50 ms. Pipe-driving pulse duration, defined as the time interval between the arrival of 5% and 95% of the total pulse energy, was 0.1–0.3 s underwater and 0.2–0.6 s in air. As this exceeds 50 ms, one might conclude that the pulses were sufficiently long that the threshold data plotted in Figs. 5 and 7 (mainly based on test signals 0.5 s in duration) would apply directly.

In June most of the seals were basking on top of the ice and any remaining snow, whereas in July many were in the water as the shorefast ice was broken up and only scattered floes remained. However, if a ringed seal is hauled out on the ice under the snow (i.e., in a lair), as is common prior to June, the in-air SPLs reaching its ear will be dampened by the snow. Blix and Lentfer (1992) measured in-air sounds of a helicopter taking off next to an artificial polar bear den in Prudhoe Bay. Sound-pressure levels inside the den (covered

by less than 1 m of snow) were 38 dB lower than outside the den, 3 m away. Thus, our in-air results reflect the greatest estimates of audibility and would not apply to seals when they are in lairs.

The locations closest to the island in which ringed seals were seen were in the moat (in June) or in the open water right next to the shore (in July). During both months, the shortest distance from the pipe-driving to the water was about 40 m. According to the June model [Fig. 4(a)], the underwater SPL at that distance would be ~153 dB *re*: 1 μPa, which is well below the 180 dB (rms) ''safety level'' set by NMFS (2000). In July the model did not seem appropriate for distances <200 m, since SPLs in that range of distances were lower than expected [Fig. 4(b)]. However, even an extrapolation of the far-field model yields an underwater SPL of ~160 dB *re*: 1 μPa, well below 180 dB. Consequently, no shutdowns of the pipe-driving were required to protect seals.

Temporary threshold shifts (TTS) for pinnipeds exposed to sound pulses such as the ones described in this paper have not been determined. Without such data, there is no direct basis for assessing what safety criterion is appropriate for pinnipeds. NMFS (2000) assumes that, for exposure to sequences of pulsed sound, the underwater safety level is 10 dB higher for pinnipeds than for cetaceans (i.e., 190 vs 180 dB *re*: 1 μPa, rms). However, recent studies (Kastak and Schusterman, 1998, 1999; Ketten *et al.*, 2001) suggest that the pinniped auditory system may actually have a lower tolerance to strong sounds as compared with odontocetes. In any event, exposures during the present circumstances were far below NMFS' established safety criterion.

C. Seal reactions to industrial sounds

None of the 23 ringed seals seen by marine mammal monitors stationed on Northstar Island reacted strongly to the acoustic or visual stimuli arising from construction activities on the island. Thirty-nine percent (9 of 23) showed no reaction to any activity, and many of the seals (48% or 11 of 23) appeared either indifferent or curious when exposed to pipe-driving and construction sounds. For example, a seal approached to within 3 m of the island's edge during pipe-driving (distance from pipe=46 m). Other seals swam in the moat during pipe-driving and other activities in June.

Seals in the moat may have been exposed to sound levels of up to 153–160 dB *re*: 1 μPa (rms) when they dove close to the bottom, whereas at or near the surface levels were presumably weaker because of pressure release effects (Urick, 1983; Greene and Richardson, 1988). There was no indication that seals sighted in the moat were reluctant to dive, or that they were doing so for shorter periods than ringed seals observed from vessels either underway or stationary in the Prudhoe Bay area during other industry activities (Harris *et al.*, 2001; Moulton and Lawson, 2002). The juvenile seal observed on 21 June spent more than 2 min swimming (and submerging) in the water nearest to the pipe-driving site. Other seals swimming in the moat dove and resurfaced at similar intervals during periods with and without pipe-driving, but no measures of diving intervals were made.

Seals' reactions to pipe-driving (or other construction activities) were very different from their reactions to helicopter traffic: 85% (17 of 20) showed no reaction to pipe-driving, whereas 92% (11 of 12) reacted to helicopter operations. However, these reactions to helicopters were not strong or long lasting, as only 8% (1 of 12) of seals subjected to a Bell 212 helicopter passing relatively low over them returned to the water. In a study by Born *et al.* (1999), 49% of 227 seals dove in response to a Bell 206 helicopter at 150-m altitude. All three seals that reacted to any of the activities at a ''moderate pace'' did so in response to a helicopter.

Relative to the behavior of ringed seals occurring far from human activities, the seals observed near Northstar in June and July were likely habituated to the industrial sounds and visible activities around the island. Seals resident in this area had been exposed to much construction activity for a period of several months prior to our observations (Moulton *et al.*, 2003). Northstar construction activities (double shifts, 7 days a week) began on the sea ice in November 1999 with ice road construction to the island. Ice thickening to support heavy equipment continued through March 2000. After a large hole was cut in the sea ice, gravel dumping began in early February and was complete by April 2000. Other construction activities that took place before the onset of pipe-driving included vibratory sheet pile-driving, trench dredging, welding and burying of a subsea pipeline to shore, bucket dredging of the island side slopes, and installation of island facilities. During this study, ringed seals approached the island during a variety of industrial activities, suggesting that the underwater and in-air sound levels were not strong enough to elicit avoidance reactions by these individuals.

It is possible that seals intolerant of the island activities had already left the area. However, results of other ringed seal monitoring activities during the preceding winter and spring (Williams *et al.*, 2002; Moulton *et al.*, 2003) suggest that ringed seals near Northstar were tolerant of the construction activities, and probably remained in the same general area throughout the winter and spring.

V. CONCLUSION

Based on our measurements, impact pile-driving of pipes—up to 107 cm in diameter and for the particular Northstar situation described above—did not expose ringed seals near Northstar Island to sound levels that came close to the NMFS criterion of 190 dB *re*: 1 μPa rms (NMFS, 2000). None of the 23 observed animals appeared to react adversely to acoustic or visual stimuli arising from construction activities on the island, and most appeared either indifferent or curious. Basking seals reacted only mildly to in-air sounds except when a Bell 212 helicopter passed relatively low over them. We estimate pipe-driving sounds to be audible to seals to distances of less than 3 km underwater and at least 0.5 km in air. Most likely, the ringed seals observed near Northstar were habituated to industrial sounds and activities around the island, and may have had a higher tolerance for such disturbances than seals from remote areas of the Arctic.

ACKNOWLEDGMENTS

We thank Dr. Charles Greene of Greeneridge Sciences, along with Jonah Leavitt of LGL, for making the June field recordings. Dr. Greene also provided valuable comments on the manuscript. Bob Blaylock of Greeneridge Sciences helped with the analysis. Jessy Coltrane and Jonah Leavitt conducted marine mammal observations and assisted with the acoustic recordings, and Shelly Schwenn and Craig Perham, all from LGL, helped with logistics. We thank Dr. Bill Streever of BP for his constructive review and for supporting the preparation of this paper as a peer-reviewed journal article, in the belief that results of industry-sponsored environmental studies should be shared with the broader scientific community. Jeff Huey, Earl Beverly, and Bob Hall (BP-Northstar), and Jim Workman, Randy Greenway, Mike Brady, Butch Smaker, and Don Pohland (Alaska Interstate Construction), provided LGL with valuable information and logistical support. Dr. Beez Hazen (Northern Engineering and Scientific) provided information about the impact driving system and schedule. ERA helicopter crews and the staff on the crew boat HAWK provided transportation; the crew boat HAWK was also the recording platform during the July measurements. Dr. W. John Richardson of LGL provided program direction and guidance, as well as helpful comments. We thank Dr. Ronald Schusterman for permission to use the recent harbor seal audiometric data. Finally, we thank Dr. Ray Jakubczak and Dr. Bill Streever of BP Exploration (Alaska), and Dave Trudgen of OASIS Environmental, for their support of this project, as well as two anonymous reviewers for comments that greatly improved the manuscript.

Bjorno, L. (**1977**). ''Finite-amplitude wave propagation through water-saturated marine sediments,'' Acustica **38**, 195–200.

Blix, A. S., and Lentfer, J. W. (**1992**). ''Noise and vibration levels in artificial polar bear dens as related to selected petroleum exploration and developmental activities,'' Arctic **45**(1), 20–24.

Born, E. W., Riget, F. F., Dietz, R., and Andriashek, D. (**1999**). ''Escape responses of hauled out ringed seals (*Phoca hispida*) to aircraft disturbance,'' Polar Biology **21**(3), 171–178.

Burgess, W. C., and Greene, Jr., C. R. (**1999**). ''Physical acoustics measurements,'' in *Marine Mammal and Acoustical Monitoring of Western Geophysical's Open-water Seismic Program in the Alaskan Beaufort Sea, 1998*, edited by W. J. Richardson, LGL Report TA2230-3. (Report from LGL Ltd., King City, Ont., and Greeneridge Sciences Inc., Santa Barbara, CA, for Western Geophysical, Houston, TX, and Nat. Mar. Fish. Serv., Anchorage, AK, and Silver Spring, MD), pp. 3–1–3–65.

Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay, Winnetka, IL), pp. 327–328.

Greene, Jr., C. R. (**1999**). ''Pile-driving and vessel sound measurements during installation of a gas production platform near Sable Island, Nova Scotia, during March and April, 1998,'' Report 205-2. (Report by Greeneridge Sciences Inc., Santa Barbara, CA and LGL Ltd., Environmental Research Associates, King City, Ont., for Sable Offshore Energy Project, Halifax, NS). Available from ExxonMobil Canada, PO Box 517, Halifax, NS, B3J 2R7 Canada.

Greene, Jr., C. R., and Burgess, W. C. (**2000**). ''Physical acoustics measurements, 1999,'' in *Marine Mammal and Acoustical Monitoring of Western Geophysical's Open-water Seismic Program in the Alaskan Beaufort Sea, 1999*, edited by W. J. Richardson, LGL Report TA2313-4. (Report from LGL Ltd., King City, Ont., and Greeneridge Sciences Inc., Santa Barbara, CA, for Western Geophysical, Anchorage, AK, and Nat. Mar. Fish. Serv., Anchorage, AK, and Silver Spring, MD), pp. 3–1–3–45.

Greene, Jr., C. R., and Richardson, W. J. (**1988**). ''Characteristics of marine seismic survey sounds in the Beaufort Sea,'' J. Acoust. Soc. Am. **83**(6), 2246–2254.

Hamilton, E. L. (**1980**). ''Geoacoustic modeling of the sea floor,'' J. Acoust. Soc. Am. **68**(5), 1313–1340.

Harris, F. J. (**1978**). ''On the use of windows for harmonic analysis with the discrete Fourier transform,'' Proc. IEEE **66**(1), 51–83.

Harris, R. E., Miller, G. W., and Richardson, W. J. (**2001**). ''Seal responses to airgun sounds during summer seismic surveys in the Alaskan Beaufort Sea,'' Marine Mammal Sci. **17**(4), 795–812.

Johnson, C. S. (**1968**). ''Relation between absolute threshold and duration-of-tone pulses in the bottlenosed porpoise,'' J. Acoust. Soc. Am. **44**(4), 965–967.

Johnson, C. S. (**1991**). ''Hearing thresholds for periodic 60-kHz tone pulses in the beluga whale,'' J. Acoust. Soc. Am. **89**(6), 2996–3001.

Johnson, S. R., Greene, Jr., C. R., Davis, R. A., and Richardson, W. J. (**1986**). ''Bowhead whales and underwater noise near the Sandpiper Island drillsite, Alaskan Beaufort Sea, autumn 1985,'' (Report from LGL Ltd., King City, Ont., and Greeneridge Sciences Inc., Santa Barbara, CA, for Shell Western E & P Inc., Anchorage, AK).

Kastak, D., and Schusterman, R. J. (**1998**). ''Low-frequency amphibious hearing in pinnipeds: Methods, measurements, noise, and ecology,'' J. Acoust. Soc. Am. **103**(4), 2216–2228.

Kastak, D., and Schusterman, R. J. (**1999**). ''In-air and underwater hearing sensitivity of a northern elephant seal (*Mirounga angustirostris*),'' Can. J. Zool. **77**(11), 1751–1758.

Kastak, D., Schusterman, R. J., Southall, B. L., and Reichmuth, C. J. (**1999**). ''Underwater temporary threshold shift induced by octave-band noise in three species of pinniped,'' J. Acoust. Soc. Am. **106**(2), 1142–1148.

Kelly, B. P., and Quakenbush, L. T. (**1990**). ''Spatiotemporal use of lairs by ringed seals (*Phoca hispida*),'' Can. J. Zool. **68**(12), 2503–2512.

Ketten, D. R., O'Malley, J., Moore, P. W. B., Ridgway, S., and Merigo, C. (**2001**). ''Aging, injury, disease, and noise in marine mammal ears,'' J. Acoust. Soc. Am. **110**(5), 2721.

Malme, C. I., Greene, Jr., C. R., and Davis, R. A. (**1998**). ''Comparison of radiated noise from pile-driving operations with predictions using the RAM model,'' LGL Report TA2224-2. (Report by LGL Ltd., Environmental Research Associates, King City, Ont., Engineering and Scientific Services, Hingham, MA, and Greeneridge Sciences Inc., Santa Barbara, CA for Sable Offshore Energy Project, Halifax, NS, Canada). Available from ExxonMobil Canada, PO Box 517, Halifax, NS, B3J 2R7 Canada.

McCauley, R. D., Jenner, M.-N., Jenner, C., McCabe, K. A., and Murdoch, J. (**1998**). ''The response of humpback whales (*Megaptera novaeangliae*) to offshore seismic survey noise: Preliminary results of observations about a working seismic vessel and experimental exposures,'' APPEA (Australian Petroleum Production and Exploration Association) J. **38**, 692–707.

McCauley, R. D., Fewtrell, J., Duncan, A. J., Jenner, C., Jenner, M.-N., Penrose, J. D., Prince, R. I. T., Adhitya, A., Murdoch, J., and McCabe, K. A. (**2000**). ''Marine seismic surveys—A study of environmental implications,'' APPEA (Australian Petroleum Production and Exploration Association) J. **40**, 692–708.

Møhl, B. (**1968**). ''Auditory sensitivity of the common seal in air and water,'' J. Aud Res. **8**(1), 27–38.

Moulton, V. D., and Lawson, J. W. (**2002**). ''Seals, 2001,'' in ''*Marine Mammal and Acoustical Monitoring of WesternGeco's Open-water Seismic Program in the Alaskan Beaufort Sea, 2001*,'' edited by W. J. Richardson (Report from LGL Ltd., King City, Ont., and Greeneridge Sciences Inc., Santa Barbara, CA, for Western Geophysical, Houston, TX, and Nat. Mar. Fish. Serv., Anchorage, AK, and Silver Spring, MD), LGL Report TA2564-3, pp. 3–1–3–48.

Moulton, V. D., Richardson, W. J., McDonald, T. L., Elliott, R. E., and Williams, M. T. (**2002**). ''Factors influencing local abundance and haulout behaviour of ringed seals (*Phoca hispida*) on landfast ice of the Alaskan Beaufort Sea,'' Can. J. Zool. **80**(11), 1900–1917.

Moulton, V. D., Richardson, W. J., Williams, M. T., and Blackwell, S. B. (**2003**). ''Ringed seal densities and noise near an icebound artificial island with construction and drilling,'' ARLO **4**(4), 112–117.

NMFS. (**2000**). ''Taking and importing marine mammals; taking marine mammals incidental to construction and operation of offshore oil and gas facilities in the Beaufort Sea,'' Fed. Regist. **65**(102, 25 May), 34014–34032.

Officer, C. B. (**1958**). *Introduction to the Theory of Sound Transmission with Application to the Ocean* (McGraw-Hill, New York.), p. 122.

Richardson, W. J., Greene, Jr., C. R., Malme, C. I., and Thomson, D. H. (**1995**). *Marine Mammals and Noise* (Academic, San Diego, CA), pp. 127–132, 346.

Schusterman, R. J., Kastak, D., Southall, B. L., Reichmuth Kastak, C., and

Holt, M. M. (**2003**). ''Noise-induced temporary threshold shift in pinnipeds: Effects of exposure medium, intermittence, duration, and intensity,'' Presentation at ECOUS (Environmental Consequences of Underwater Sound) meeting, 12–16 May, 2003, San Antonio, TX.

Sivian, L. J., and White, S. D. (**1933**). ''On minimum audible sound fields,'' J. Acoust. Soc. Am. **4**(4), 288–321.

Terhune, J. M. (**1988**). ''Detection thresholds of a harbour seal to repeated underwater high-frequency, short-duration sinusoidal pulses,'' Can. J. Zool. **66**(7), 1578–1582.

Terhune, J. M. (**1989**). ''Underwater click hearing thresholds of a harbour seal, *Phoca vitulina*,'' Aquatic Mammals **15**(1), 22–26.

Terhune, J. M., and Ronald, K. (**1975**). ''Underwater hearing sensitivity of two ringed seals (*Pusa hispida*),'' Can. J. Zool. **53**(3), 227–231.

Terhune, J., and Turnbull, S. (**1995**). ''Variation in the psychometric functions and hearing thresholds of a harbour seal,'' in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. W. Nachtigall (De Spil., Woerden, Netherlands), pp. 81–93.

Urick, R. J. (**1983**). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York, Reprinted 1996, Peninsula, Los Altos, CA), pp. 131–134.

Williams, M. T., Smith, T. G., and Perham, C. J. (**2002**). ''Ringed seal structures in sea ice near Northstar, winter and spring of 2000–2001,'' in *Monitoring of Industrial Sounds, Seals, and Whale Calls during Construction of BP's Northstar Oil Development, Alaskan Beaufort Sea, 2001*, edited by W. J. Richardson and M. T. Williams [Oct. 2002 ed.] LGL Rep. P557-2. (Report from LGL Ltd., King City, Ont., and Greeneridge Sciences Inc., Santa Barbara, CA, for BP Explor. (Alaska) Inc., Anchorage, AK, and Nat. Mar. Fish. Serv., Anchorage, AK, and Silver Spring, MD), pp. 4–1–4–33.

Würsig, B., Greene, Jr., C. R., and Jefferson, T. A. (**2000**). ''Development of an air bubble curtain to reduce underwater noise of percussive piling,'' Marine Environmental Research **49**, 79–93.

The influence of viscosity on the shear strain remotely induced by focused ultrasound in viscoelastic media

E. A. Barannik, S. A. Girnyk, and V. V. Tovstiak
Kharkiv National University, 4 Svobody Sq., Kharkiv, Ukraine 61077

A. I. Marusenko and V. A. Volokhov
JSC Research Development Institute of Radio Engineering Measurements, 271 Ac. Pavlov Av., Kharkiv, Ukraine 61054

A. P. Sarvazyan[a)]
Artann Laboratories, 1753 Linvale-Harbourton Road, Lambertville, New Jersey 08530

S. Y. Emelianov
Department of Biomedical Engineering, University of Texas at Austin, 2501 Speedway, ENS 636, Austin, TX 78712

(Received 21 February 2003; revised 30 January 2004; accepted 10 February 2004)

Shear wave elasticity imaging (SWEI), an emerging acoustic technology for medical diagnostics, is based on remote generation of shear waves in tissue by radiation force in the focal region of an ultrasonic beam. In this study, the feasibility of Doppler ultrasonic technique to visualize the remotely induced shear waves was demonstrated. The generation of shear displacement in the focal region of a pulsed 1-MHz ultrasound beam with pulse duration of approximately about 2 ms and intensity levels on the order of 145 W/cm^2, and consequent propagation of shear wave in tissue-mimicking and muscle tissue *in vitro*, were measured. The analysis of temporal behavior of shear displacement within the focal plane allowed estimation of shear wave velocities. The velocities were 4 and 7 m/s in hard phantom and tissue containing phantom, respectively. The measured shear displacements on the order of micrometers in gel-based phantoms are in reasonable agreement with theoretical estimates derived from an earlier developed model of shear wave generation by radiation force of focused ultrasound. The study revealed significant dependence of shear strain on the medium viscosity. The complex oscillatory character of shear strain relaxation in viscoelastic phantom and muscle tissue *in vitro* was observed. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1698796]

PACS numbers: 43.80.Qf, 43.80.Sh, 43.80.Vj [FD] Pages: 2358–2364

I. INTRODUCTION

The feasibility of shear wave elasticity imaging (SWEI) was previously demonstrated in experiments with optical and NMR detection of ultrasonically induced shear waves.[1–3] Propagation of shear waves was also demonstrated using an ultrafast ultrasound imaging system.[4] SWEI may be considered as a new branch of ultrasound elasticity imaging that is being extensively developed and investigated in a number of laboratories.[4–15]

The core of the SWEI method is generation of shear strain in tissue using radiation force of focused ultrasound, as it is schematically shown at Fig. 1. The remotely induced shear strain can be detected using various ultrasonic or MRI techniques,[1,12–15] after which the obtained ''palpation information'' is analyzed and the tissue viscoelastic properties are evaluated. The diagnostic value of tissue mechanical properties depends on the range of variation of these characteristics prompted by changes in tissue condition. The mechanical attributes of tissue sensed during palpation reflect its shear elastic modulus—a parameter that varies over several orders of magnitude.[16]

At present, several laboratories conduct investigations of various radiation force based imaging methods.[1,4,12–15] All proposed methods include the remote generation of an elastic deformation in tissue using a focused ultrasonic beam and a subsequent noninvasive monitoring of tissue response. These methods differ in design, implementation, and tissue response interpretation. In our approach and realization of SWEI the radiation force or excitation beam is fired in a single direction. Figure 1 shows the shear strain induced by the radiation force in the focal plane of the ultrasound excitation transducer and the resulted shear wave propagating radially from the beam axis.

The shear displacement induced at the beam axis by a short pulse of the radiation force of a circular focused ultrasonic beam with a Gaussian amplitude profile in the focal plane can be presented in the following form:[17]

$$S(r=0,t)=\frac{\sqrt{\pi}}{2}\frac{\alpha(aD)^2 I}{\rho c_t c}\exp(-2\alpha d)\times\left(\frac{t_0}{aD}\right)\frac{(c_t/aD)t}{1+4\nu t/(aD)^2+(c_t/aD)^2 t^2}, \quad (1)$$

for transducers with relatively high *f*-numbers, i.e., where

[a)]Author to whom correspondence should be addressed. Electronic mail: armen@artannlabs.com

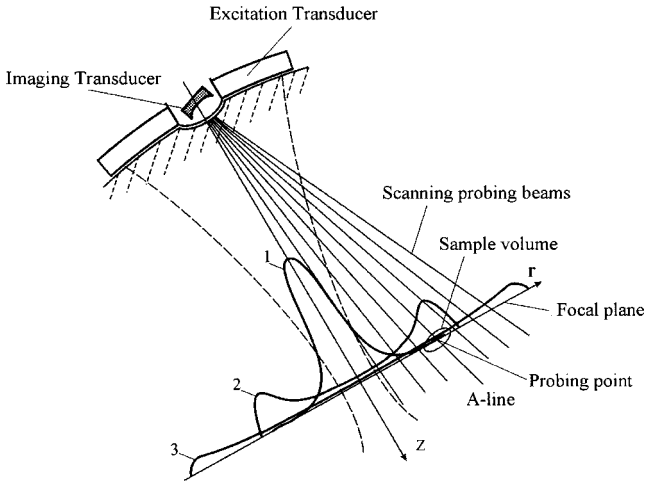


FIG. 1. Schematic presentation of ultrasound transducer assembly used in performed SWEI experiments. The excitation transducer produces a shear deformation in the focal region with initial profile (1), and the imaging transducer, incorporated in the middle of and aligned with the excitation transducer, that can be used both to image the object in conventional B-mode and to detect the propagation of shear waves (2 and later 3) in the focal plane of excitation beam.

axial size of the focal region is much larger than the transverse. The displacement reaches maximum

$$S_{\max}=\frac{\sqrt{\pi}}{4}\,\frac{\alpha(aD)I\cdot t_0}{\rho c_t c}\left(1+\frac{2\eta}{\rho c_t aD}\right)^{-1}\exp(-2\alpha d), \tag{2}$$

at $t_{\max}=aD/c_t$. Here, a is the radius of Gaussian pumping beam at a e^{-1} level relative to that at the transducer surface so that the effective aperture W is equal to $2a$, d is the curvature radius of the transducer active element given by the boundary conditions on its surface, t_0 is the duration of excitation pulses, I is the spatial peak pulse average intensity of the excitation beams at the beam axis in the focal plane, c and c_t are the speeds of ultrasound (compressional) and shear waves in the medium, respectively, η and $\nu=\eta/\rho$ are dynamic and kinematic viscosity, respectively, and α is the ultrasonic absorption coefficient for longitudinal waves at the given frequency in the tissue. The parameter D in (1) and (2) is inversely proportional to the degree of focusing: $D^{-1}=\gamma=l_F/d$, where $l_F=ka^2/2=\pi a^2/\lambda$ is the Fresnel length[18] for circular transducer of radius a, k is a wave number, and λ is the ultrasound wavelength. Thus, the value of initial displacement is inversely proportional to the velocity c_t that is directly related to the shear elasticity modulus with a well-known ratio: $c_t=\sqrt{\mu/\rho}$, where μ is the shear modulus.

The expression (1) accounts for the viscosity of the medium and can be used to assess the role of viscosity in shear response induced by radiation force of the focused ultrasonic beam. Currently, the influence of viscosity on displacement in the focal region has neither been carefully studied nor discussed in literature. However, soft tissues are viscoelastic, and it is important to gather an insight into possible influence of media viscosity on the propagation parameters of the radiation force induced shear waves. Thus, the objectives of present study were to detect shear waves propagating radially from the focal area of a pulsed ultrasonic beam in tissue-mimicking phantoms and soft tissue *in vitro*, and to explore the influence of viscosity on shear strain induced by radiation force of focused ultrasonic beam.

TABLE I. Tissue-mimicking phantoms.

	Phantom A	Phantom B
Water, liter	1.1	2.5
Glycerol, liter	1.26	
Gelatin, g/liter	58	96
Speed of shear waves, c_t (20 °C)	2.3 m/s	3.6 m/s
Speed of ultrasound, c (20 °C)	1810 m/s	1560 m/s
Ultrasound absorption	0.36 dB/MHz.cm	0.51 dB/MHz.cm

II. METHODS AND MATERIALS

The shear wave elasticity imaging experiments were performed using specifically designed tissue-mimicking and tissue-containing phantoms. Two sets of phantoms were produced. In the first set, phantoms were constructed from different gel concentrations. Initially, gelatin was mixed with water at room-temperature and kept for 1 h. Then the mixture was heated up to 80 °C until the gelatin was completely dissolved, and chilled to 20 °C in about an hour. In addition, once the gelatin was dissolved, a glycerol—a viscous, water-solvable liquid—was added into phantom A to increase the shear viscosity of the material. The properties of both gel-based phantoms (phantom A and phantom B) are summarized in Table I. Per phantom composition, phantom A has a greater shear viscosity due to the glycerol additive while phantom B has a greater shear modulus due to an increased concentration of gelatin. In comparison with the ''viscous'' phantom A, the ''hard'' phantom B has a greater ultrasound attenuation that was measured by the common transmission method[19] using the short ultrasound bursts and 60-mm-thick samples. This apparent incongruity (i.e., decreased viscosity of material in phantom B) can be explained by the fact that the absorption of a compressional wave, contrary to shear, does not depend purely on macromolecular conformation relaxation processes.[20] Therefore, the elevated concentration of gelatin in phantom B resulted in an increased attenuation of ultrasound longitudinal waves.

The second set of phantoms had embedded soft tissue samples. Initially, the large specimen of bovine muscle was cut into two nearly identical samples. The first sample was cooked in boiling water for 1 h, while the second sample was kept on ice. Then each sample was positioned into a thin rubber sack serving as mold, and the gelatin solution at temperature 30 °C was added into the mold to create the final ''hard'' and ''viscous'' phantoms. Both samples were positioned such that the orientation of fibers was transverse relative to the direction of the excitation beam and the imaging plane within which the propagation of shear waves were measured. Since the objective of the current study was to reveal qualitative differences in the ''viscous'' versus ''non-viscous'' objects, we did not attempt to characterize quantitatively the value of the dynamic shear viscosity of the samples.

The shear waves in phantoms were generated using 1 MHz, 80-mm-diam single element, focused (f-number=1)

TABLE II. Parameters of shear wave elacticity imaging system.

	Excitation	Probing
Focal area diameter (−6 dB)	2.0 mm	1.8 mm
Focal length	60 mm	80 mm
Carrier frequency	1 MHz	3.5 MHz
Pulse duration t_0	100 μs−10 ms	1.7 μs
PRF	14.59 Hz	3.67 kHz
I_{SPPA} in distilled water	145 W/cm^2	
I_{SPTA} in distilled water	4.25 W/cm^2 (t_0=2.18 ms)	40 mW/cm^2

excitation transducer with an opening in the middle for an imaging transducer, to allow both imaging of phantoms in B-mode and ultrasonic Doppler detection of induced shear waves (Fig. 1). The overall parameters of both transducers and the corresponding excitation and probing beams of waves and excitation pulses are given in Table II. Both the imaging and excitation transducers were assembled into one unit and interfaced with the commercial diagnostic scanner TI628 (JSC Research Development Institute of Radio Engineering Measurements, Ukraine). The scanner has been modified to allow a user to control several critical settings of the experiment including the duration, intensity, and pulse repetition frequency (PRF) of excitation pulses, the interleaving factor between probing and excitation beams, and the volume of the range cell (or sample volume) formed by probing pulses. The excitation and probing pulses were interspersed and the number of probing pulses that follow every excitation pulse was chosen prior to the experiment.

On-line data processing was performed using the phase-sensitive method similar to one used in conventional pulsed Doppler and color flow systems to measure the axial component of blood flow. For an isolated scatterer positioned along the beam, the basebanded representation of ultrasound signal can be written as

$$s_1(t)=s(t-\tau_1)=A(t-\tau_1)e^{-j\omega_0\tau_1},$$

where ω_0 is the angular frequency of the ultrasound carrier, τ_1 is the round trip propagation time from the transducer to the point scatterer and A is the envelope function of the pulse. Using this relationship, the position of the scatterer $x_1(t)$ is related to the phase of the ultrasound signal

$$x_1(t)=\frac{\lambda}{4\pi}\varphi(t)=\frac{\lambda}{4\pi}\tan^{-1}\frac{\mathrm{Im}(s_1)}{\mathrm{Re}(s_1)},$$

where phase $\varphi(t)$ is unwrapped and defined within $(-\pi,\pi)$ range. Assuming that the radiation force induced axial motion between two probing firings is such that Nyquist condition

$$|\varphi_n(t)-\varphi_{n-1}(t)|<\pi$$

is satisfied, the displacement at each point can be measured based on a phase change in ultrasound signals $s_n(t)$ and $s_{n+1}(t)$.

To overcome the effect of interference between the excitation and probing beams (due to the power level of the excitation pulse and associated backscattering of excitation waves received during probing), a pulse inversion technique was used. Specifically, every even excitation pulse was inverted respectively to the odd pulse. For each pulse, two ultrasound Doppler signals, $s^0(t)$ and $s^\pi(t)$, were recorded where the first signal was obtained after the odd excitation pulse, and the second signal was obtained after the inverted, even excitation pulse. Adding two signals together, $s(t)=s^0(t)+s^\pi(t)$, the residual excitation signal was significantly minimized. In addition, the input radio-frequency signals were accumulated coherently, thus increasing the signal-to-noise ratio (SNR) of the displacement measurements.

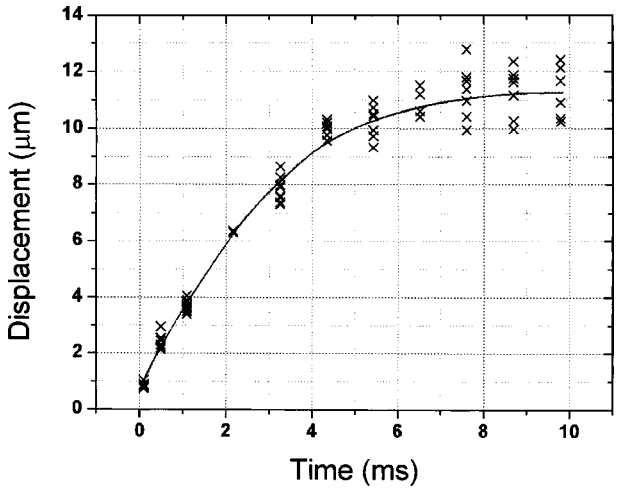


FIG. 2. The maximum displacement amplitudes produced by single excitation pulse of varying duration. The measurements were performed using phantom A. For each pulse duration, several measures were taken (denoted by x). A solid line represents least-squares polynomial fit of the measured data.

The results of the experiments reported here were obtained using t_0=2.18-ms duration excitation pulse with 14.59-Hz pulse repetition frequency (PRF). In these experiments, the described excitation of shear displacement and strain should result in a sufficiently low, about 4.5 W/cm^2, level of intensity I_{SPTA} which is within the range of diagnostic intensities of Doppler ultrasound.[21] After each excitation pulse, a large, on the order of 250, number of probing pulses were transmitted to allow the Doppler-based registration of the induced shear strain and subsequent stress relaxation, while the coherent data accumulation was performed for each pair of excitation pulses. As a result, the standard deviation in shear displacement measurements was within 5%–10% of the magnitude.

III. RESULTS

First, a set of experiments was performed to estimate the optimum duration of the excitation pulse. Figure 2 presents a relationship between the duration of the excitation pulse and resulting displacement. The displacements were measured in phantom A with the excitation and probing beams aligned with each other. The pulse duration of about 2 ms appears to be a reasonable trade-off between efficient generation of shear strain in this phantom and minimized delivery of sound, and, therefore, nonlinear effects and thermal energy discharged inside the focal region.

Figures 3 and 4 present the relaxation dynamics of the local displacement induced in ''hard'' phantom B and ''vis-

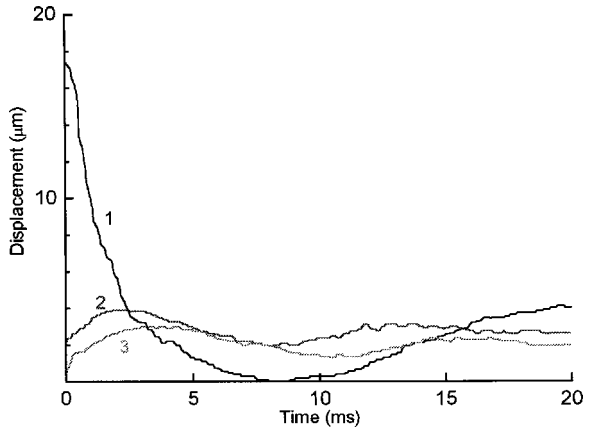


FIG. 3. Temporal behavior of the shear displacement in "hard" phantom B recorded at various locations within the focal plane (see Fig. 1). The 2.18-ms excitation pulse was used to produce the shear displacement in the focal region centered at $r=0$ mm, and the radiation force induced displacements were measured at (1) $r=0$ mm, (2) $r=11$ mm, and (3) $r=15$ mm.

cous" phantom A, correspondingly. The curves labeled with "1" in both figures depict the temporal behavior of the axial displacement at the focal point where the axes of the excitation and probing beams were aligned—the radial coordinate on the probing beam at this position is arbitrarily set to zero. The propagation of the shear wave packet in the lateral direction in phantom B is demonstrated at additional two points with radial positions of 11 mm (curve 2) and 15 mm (curve 3). For the phantom A, the propagation of shear waves is illustrated using measurements at the point with radial position of 11 mm. The time delay of the wavefront arrival to the given point increases as the radial coordinate of the point increases. Clearly, the group velocity of the shear wave packet, depending in particular on the shear elasticity, can be calculated from the measurements of wavefront propagation. Detailed time-of-flight analysis of curves 1, 2, and 3 presented in Fig. 3 reveals that the shear wave speed in hard phantom B is approximately 4 m/s. The value of the shear wave propagation speed in homogeneous objects, generally defined as ratio of propagation distance (r) and time required (Δt), $c_t=r/\Delta t$, can be determined more accurately by measuring the delay time Δt at several points with different radial coordinates r and by either averaging or using a first-order polynomial fit of the obtained data as shown in Fig. 5. These averaged values of the shear wave speed in both phantoms are given in Table I. As the shear viscosity and, therefore, attenuation of shear waves increases, the peak of shear displacement is less pronounced and cannot be accurately identified while shear wave propagates outward from the center. In viscous phantom A, for example, the maximum cannot be reliably identified at a distance greater than 15 mm.

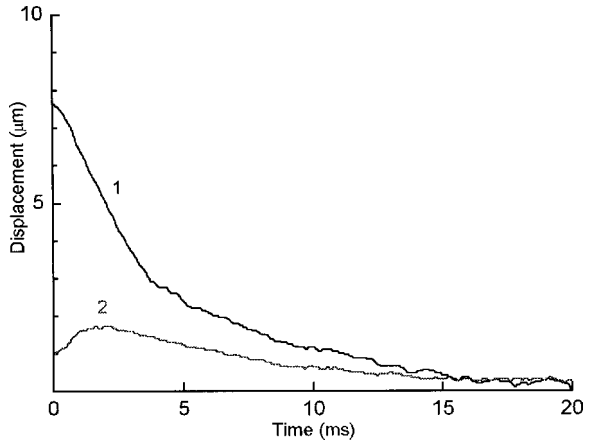


FIG. 4. Temporal behavior of the shear displacement generated in "viscous" phantom A using 2.18-ms excitation pulse. The radiation force induced displacements were measured at (1) $r=0$ mm, and (2) $r=11$ mm.

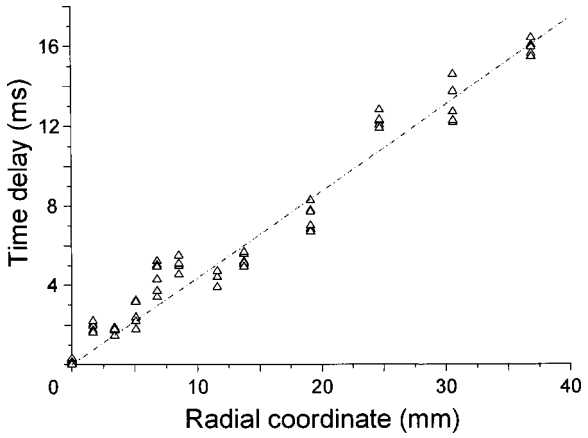


FIG. 5. Propagation of shear wave packet in viscous phantom. The shear displacements were induced by radiation force of 2.18-ms excitation pulse of focused ultrasound at $r=0$ mm, and measured at several radial positions as indicated on the plot. A line represents least-squares polynomial fit of the measured data, from which the average speed of a shear wave packet of 2.3 m/s was determined and presented in Table I.

Figure 6 presents the measurement results for the tissue-containing phantoms. The curves 1 and 3 in this figure are temporal behaviors of the axial displacements at the focal point of the excitation transducer of "hard" (i.e., boiled tissue) and "viscous" (i.e., raw tissue) phantoms, correspondingly. The curves 2 and 4 are the axial displacements captured 30 and 20 mm away from the focal area of the excitation transducer in hard and viscous phantoms, respectively. As expected, the speed of shear waves in the boiled

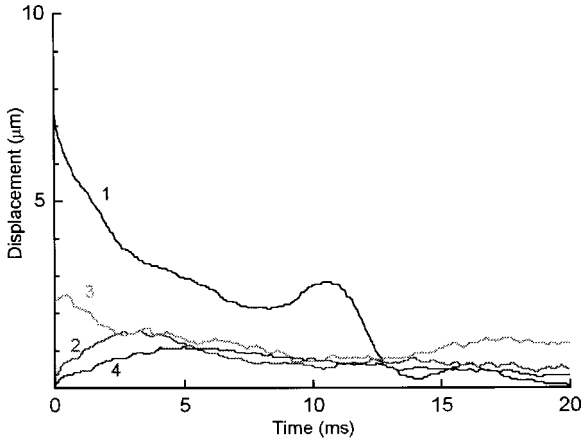


FIG. 6. Comparison of shear displacement relaxation in gel-based phantoms with embedded boiled (lines 1 and 2) and raw (lines 3 and 4) beef muscle sample. In all measurements, the shear displacements were produced using 2.18-ms excitation pulse. The radiation force induced displacements were measured at $r=0$ mm (1 and 3), $r=20$ mm (4), and $r=30$ mm (2).

tissue is substantially higher than that in the raw muscle due to an increased shear modulus. Comparing curves 1 and 2 in Fig. 6, the shear wave velocity in this tissue is about 7 m/s. Note that the relatively high speed of shear wave propagation in the boiled muscle complicates time delay analysis of the displacement relaxation curves near the excitation point.

IV. DISCUSSION

Initially, the magnitude of the shear displacement generated by the focused ultrasound beam is almost linearly related to the duration of the ultrasound pulse (Fig. 2). This relationship is in agreement with the expression (1) derived by Rudenko *et al.*[17] under certain simplifying assumptions. This expression adequately describes the excitation of shear strain produced in elastic media and biological tissues using relatively short ultrasound pulse. In fact, it is possible to use the time needed for the displacement to reach the maximum as a tissue characterization parameter. However, with the increase of the pulse duration, the displacement magnitude approaches to a certain saturation determined primarily by the elastic properties of the medium. The limit case of long excitation pulses ($t=aD/c_t>1$ ms) corresponds to localized static loading that, for example, was discussed in Ref. 22. Evidently, the saturation value can also be used as a tissue characterization parameter, although there is an increased level of ultrasonic intensity due to prolonged excitation pulses needed to evaluate the saturation.

The difference in shear elastic moduli between phantoms A and B, proportional mainly to gelatin concentration, resulted in difference of shear wave speed. This is clearly demonstrated in Figs. 3 and 4, where curve 2 for both figures was obtained 11 mm away from the origin of shear strain excitation. The wavefront and maximum of the shear displacement expectedly reached this point earlier for phantom B than compared to phantom A. The displacement amplitude in both phantoms significantly decreases with the propagation distance (Figs. 3 and 4). This is a result of both the absorption of high-frequency wave components of the induced shear strain, and the geometrical divergence of cylindrical shear waves formed in the focal area and propagating outward from the excitation site.

In spite of the lower shear modulus of phantom A, the amplitude of induced displacements is substantially lower than that in phantom B (Figs. 3 and 4). The amplitude of the radiation force and, correspondingly, the magnitude of shear displacement (2) are proportional to the ultrasound absorption coefficient and magnitude of ultrasound signal in the focal region. Clearly, the ultrasound signal attenuates as it reaches the focal area—this effect is described by the exponential term in (1) and (2). These terms are approximately equal in viscous (A) and hard (B) phantoms, although the same 2.18-ms radiation force pulse results in a lesser displacement amplitude in the phantom A due to the higher viscosity (Table I), as it was predicted by (2).

The temporal relaxation of the shear displacement is also very different for the two phantoms used in the first set of experiments. Due to the elevated viscosity in phantom A, the shear displacement, once it reaches the maximum, relaxes more slowly compared to that of phantom B. In particular, in phantom A the shear displacement value decreases by two times within approximately 2.5 ms while in phantom B it takes approximately 1 ms. This observation, somewhat analogous to creep, clearly suggests that the material in phantom A is more viscous compared to material in phantom B. From our experimental results, contrary to (1), the maximum displacement in the focal region of the viscous phantom A is not always smaller compared to the elastic phantom B. This can be due to several experimental factors and further investigation of the phenomenon using phantoms with independently known elasticity and viscosity is needed.

For a given duration and intensity of excitation pulses, geometrical characteristics of the ultrasound excitation transducer and resulting ultrasound field, and physical characteristics of the phantoms, the measured shear displacements are in a reasonable agreement with the available theoretical evaluations. In expressions (1) and (2), derived for a radiation force produced by a Gaussian focused ultrasound beam,[1] the physical meaning of the $2aD$ product corresponds to the beam width in the focal region at the e^{-1} level, or approximately -8.7 dB. For the excitation transducer used in this study, $2aD=3.1$ mm is bigger than the 2.0-mm width of the excitation beam measured at -6 dB (Table II). The intensity I of excitation ultrasound pulse in (1) corresponds to the I_{SPPA} intensity. Therefore, omitting the viscosity in (2), the maximum shear displacement $S_{\max}$ is 2.1 μm for $\alpha=0.041$ cm^{-1}, $I=I_{SPPA}=145\ \mathrm{W/cm}^2$, $t_0=2.18$ ms, $\rho=1\ \mathrm{g/cm}^3$, $aD=1.55$ mm, $d=6$ cm, $c_t=2.3$ m/s, and $c=1810$ m/s.

The estimated theoretical value of $S_{\max}$ is smaller than those obtained experimentally. This is due to several assumptions used in the derivation (1) and (2). In particular, the assumed transverse distribution of the acoustic pressure amplitude within the beam may not be valid for the experimental system used in this study. Indeed, the distribution of acoustic pressure near the surface of the excitation transducer with a hole in the middle (to accommodate the imaging transducer) differs substantially from the Gaussian distribution used in theoretical analysis.[1] In addition, the ultrasound attenuation coefficient of the material was measured using bursts of low intensity ($I_{SPPA}=50\ \mathrm{mW/cm}^2$) plane waves. The nonlinear interaction of ultrasound and medium[23] further increases attenuation and absorption of ultrasound and, subsequently, the magnitude of the radiation force—the nonlinear effects were not considered in our study.

Based on the model prediction (1), the shear displacement in the focal region increases rapidly, reaches its maximum (2), and then, gradually, monotonically decreases. The measured temporal behavior of shear displacement in elastic phantom B and phantom with boiled bovine muscle does not generally agree with the model (1). The shear displacement in phantom B, for example, exhibits fluctuation in time (Fig. 3) suggesting that a relaxation process of oscillatory nature is taking place. Indeed, this phenomenon can be one of the reasons contributing to the discrepancy between the measured and theoretically predicted amplitudes of shear displacement.

The oscillatory nature of the shear displacement relaxation can be a result of several factors. First, the previously developed theory[1] assumes an infinite medium with uniform

distribution of shear elastic modulus, viscosity, and ultrasound absorption coefficient. Both the tissue-containing phantoms and tissue-mimicking phantoms satisfy these assumptions only to a certain degree. While the absorption of shear waves is only relevant over the extent of the excitation beam, the mechanical properties of the material will affect both temporal and spatial behaviors of the shear displacement. In general, analytical derivation of shear displacement and shear wave propagation in the inhomogeneous medium is a complicated problem. Nevertheless, simple physical reasoning suggests that if the characteristic size of an inhomogeneity is comparable with the size of the shear strain disturbance, the strain relaxation process can be of complex oscillatory nature due to superposition of diverging cylindrical waves and waves reflected from boundaries of local inhomogeneities. In turn, it can lead, for example, to the formation of local standing wave components of potentially large amplitude. Second, weak oscillations observed during the shear displacement relaxation can be attributed to the reflection of shear waves from the external boundaries of tissue-mimicking phantoms or gel-tissue boundaries. For a single pulse SWEI, the representative wavelength of the remotely generated shear waves is primarily determined by the geometry of the excitation beam. In our study, the 3.1-mm focal beamwidth of the excitation transducer results in approximately 6–8-mm minimum wavelength of induced shear wave packet with spectral maximum at zero frequency. Given these central frequency and bandwidth, shear waves can easily get to and reflect from the external boundaries of the phantoms or tissue samples having the dimensions about 5–10 cm. The difference in the reflection of a transverse shear wave from a rigid inhomogeneity is that the reflected shear wave has a π (180°) phase shift. Third, the expression (1) was derived for the focused Gaussian ultrasound excitation beam with no side lobes. However, a ring-shaped ultrasound transducer (i.e., transducer with a hole in the middle) is characterized by a smaller focal region in the axial direction, and by a high level of side lobes compared to a single element focused transducer of the same diameter. It is possible that the concentric rings of diverging and converging shear waves were formed in the focal plane of this transducer. Clearly, the temporal behavior of shear displacement at the focal region can be affected by these additional waves. Finally, an additional absorption of a high intensity ultrasound due to a nonlinear acoustic phenomenon[1] must be included into the model to accurately describe the interaction of ultrasound with tissue.

Measuring delay of shear displacement propagation, the shear modulus can be accurately estimated even in the presence of various internal and external boundaries and other inhomogeneities in biological tissue. Indeed, the reflected shear waves have larger delay and appear after the initial wave front (see Figs. 3 and 6, for example). These waves are also associated with smaller magnitude due to larger travel distance and, therefore, additional attenuation. Clearly, the temporal behavior of the shear displacement in viscous phantoms does not exhibit the oscillatory nature suggesting significant attenuation of shear waves. In addition to propagation delay method, local shear viscosity and elasticity can be assessed from the measurements of the displacement in the focal region.[1]

V. CONCLUSION

Results of this work confirm the findings of previous studies[1,4,13–15] that SWEI provides remote assessment of tissue mechanical properties and may add a new and desired diagnostic quality to conventional ultrasonic imaging. Present and previous[12] studies demonstrate that SWEI can be implemented using the ultrasound Doppler method for shear wave detection. Doppler-based SWEI can be implemented as a real time imaging technique. The applicability of the Doppler-based method to measure internal tissue displacement and strains induced by the pulsed radiation force of focused ultrasound was shown for a relatively low power of ultrasound excitation pulses I_{SPPA} and radiation intensity I_{SPTA}. This is critically important for Doppler-based SWEI to become a clinical tool of medical diagnostics.

The results of this study suggest that qualitative analysis of the tissue viscosity is possible in SWEI. It has been shown that the displacement relaxation in the viscous phantom requires longer time in comparison with the elastic one. Results from tissue-containing phantoms also confirm this finding. For a short duration of excitation pulses (i.e., when pulse duration is much smaller than propagation of shear wave through the focal region), the amplitude of shear displacements depends not only on shear elastic modulus but also on the viscosity. Thus, in general, the remotely induced shear displacement is an integral characteristic reflecting both elastic and viscous properties of the medium. The maximum values of shear displacements measured in gelatin-based tissue phantoms are consistent with an estimate based on the earlier developed model of shear wave excitation using radiation force phenomena.[1] The most important question to be considered in future quantitative studies is the detailed dependence of shear strain, its relaxation, and group velocity of induced wave packet on the shear viscosity and elasticity.

Possible physical factors responsible for the nonmonotonic and oscillating nature of displacement relaxation in elastic mediums observed in this study include tissue inhomogeneity, local (internal) and global (external) boundary conditions, and the diffraction pattern of the excitation transducer. Since biological tissues can be very heterogeneous, further detailed study of these factors is required to fully understand the possible temporal and spatial behavior of the shear displacement needed for shear wave elasticity imaging.

ACKNOWLEDGMENTS

The authors wish to express deep gratitude to Professor M. O'Donnell for his useful comments and fruitful discussions. This work is supported by the Science and Technology Center in Ukraine under Grant No. 865 and NIH Grant No. 1 R21 EB001548-01.

[1] A. P. Sarvazyan, O. V. Rudenko, S. D. Swanson, J. B. Fowlkes, and S. Y. Emelianov, "Shear Wave Elasticity Imaging: a new ultrasonic technology of medical diagnostics," Ultrasound Med. Biol. **24**(9), 1419–1435 (1998).
[2] V. G. Andreev, V. N. Dmitriev, Y. A. Pischalnikov, O. V. Rudenko, O. A. Sapozhnikov, and A. P. Sarvazyan, "Observation of shear waves excited

by focused ultrasound in a rubber-like media,'' Acoust. Phys. **43**, 123–128 (1997).
[3] V. G. Andreev, V. N. Dmitriev, O. V. Rudenko, and A. P. Sarvazyan, ''Remote generation of shear wave in soft tissue by pulsed radiation pressure,'' J. Acoust. Soc. Am. **102**, 3155–3158 (1997).
[4] M. Fink, L. Sandrin, M. Tanter, S. Catheline, S. Chaffai, J. Bercoff, and J. Gennisson, ''Ultra-high Speed Imaging of Elasticity,'' in Proceedings of the 2002 IEEE Ultrasonics Symposium, in print (2002).
[5] T. A. Krouskop, D. R. Dougherty, and S. F. Levinson, ''A pulsed Doppler ultrasonic system for making non-invasive measurements of the mechanical properties of soft tissue,'' J. Rehabil. Res. Dev. **24**(2), 1–8 (1987).
[6] R. M. Lerner, S. R. Huang, and K. J. Parker, ''Sonoelasticity images derived from ultrasound signals in mechanically vibrated tissues,'' Ultrasound Med. Biol. **16**(3), 231–239 (1990).
[7] K. J. Parker, S. R. Huang, R. A. Musulin, and R. M. Lerner, ''Tissue response to mechanical vibrations for 'sonoelasticity imaging,''' Ultrasound Med. Biol. **16**(3), 241–246 (1990).
[8] Y. Yamakoshi, J. Sato, and T. Sato, ''Ultrasonic imaging of the internal vibration of soft tissue under forced vibration,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **37**, 45–53 (1990).
[9] J. Ophir, I. Cespedes, H. L. Ponnekanti, Y. Yazdi, and X. Li, ''Elastography: A quantitative method for imaging the elasticity of biological tissues,'' Ultrason. Imaging **13**, 111–117 (1991).
[10] M. O'Donnell, A. R. Skovoroda, B. M. Shapo, and S. Y. Emelianov, ''Internal displacement and strain imaging using ultrasonic speckle tracking,'' IEEE Trans. Ultrason. Ferroelectr. Freq. Control **41**, 314–325 (1994).
[11] S. Y. Emelianov, A. R. Skovoroda, M. A. Lubinski, and M. O'Donnell, ''Reconstructive elasticity imaging,'' Acoust. Imaging **21**, 241–252 (1995).
[12] E. A. Barannik, S. A. Girnyk, V. V. Tovstiak, A. I. Marusenko, S. Y. Emelianov, and A. P. Sarvazyan, ''Doppler ultrasound detection of shear waves remotely induced in tissue phantoms and tissue in vitro,'' Ultrasonics **40**, 849–852 (2002).
[13] W. Walker, ''Internal deformation of a uniform elastic solid by acoustic radiation force,'' J. Acoust. Soc. Am. **105**(4), 2508–2512 (1999).
[14] K. Nightingale, R. Nightingale, M. Palmery, and G. Trahey, ''A finite element model of remote palpation of breast lesions using radiation force: factors affecting tissue displacement,'' Ultrason. Imaging **22**(1), 35–54 (2000).
[15] M. Fatemi and J. Greenleaf, ''Probing the dynamics of tissue at low frequencies with the radiation force of ultrasound,'' Phys. Med. Biol. **45**(6), 1449–1464 (2000).
[16] A. P. Sarvazyan, A. R. Skovoroda, S. Y. Emelianov, J. B. Fowlkes, J. G. Pipe, R. S. Adler, R. B. Buxton, and P. L. Carson, ''Biophysical bases of elasticity imaging,'' in *Acoustical Imaging 21*, edited by J. P. Jones (Plenum, New York, 1995), pp. 223–240.
[17] O. V. Rudenko, A. P. Sarvazyan, and S. Y. Emelianov, ''Acoustic radiation force and streaming induced by focused nonlinear ultrasound in a dissipative medium,'' J. Acoust. Soc. Am. **99**, 2791–2798 (1996).
[18] B. Novikov, O. Rudenko, and V. Timoshenko, *Nonlinear Underwater Acoustics* (American Institute of Physics, New York, 1987).
[19] J. Krautkramer and H. Krautkramer, *Ultrasonic Testing of Materials* (Springer Verlag, New York, 1983).
[20] F. W. Kremkau and R. W. Cowgill, ''Biomolecular absorption of ultrasound. I: Molecular weight,'' J. Acoust. Soc. Am. **76**, 1330–1335 (1984).
[21] F. A. Duck and J. Henderson, ''Acoustic output of modern ultrasound equipment: it is increasing?'' in *Safety of Diagnostic Ultrasound: Progress in Obstetric and Gynecological Sonography Series*, edited by S. B. Barnett and G. Kossoff (Parthenon, New York, 1998), pp. 15–24.
[22] K. Nightingale, M. Palmery, R. Nightingale, and G. Trahey, ''On the feasibility of remote palpation using acoustic radiation force,'' J. Acoust. Soc. Am. **110**, 625–634 (2001).
[23] F. A. Duck, ''Nonlinear acoustics in diagnostic ultrasound,'' Ultrasound Med. Biol. **28**(1), 1–18 (2002).

PROGRAM OF

The 147th Meeting of the Acoustical Society of America

New York, New York • Sheraton New York Hotel and Towers • 24–28 May 2004

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, 24 MAY 2004 ROYAL BALLROOM A, 7:40 A.M. TO 12:00 NOON

Session 1aAA

Architectural Acoustics and Noise: Effect of Room Acoustic Environment on Human Productivity and Performance

Rendell F. Torres, Cochair

Architectural Acoustics Program, Rensselaer Polytechnic Institute School of Architecture, 110 8th Street, Troy, New York 12180-3590

Lily M. Wang, Cochair

Architectural Engineering, University of Nebraska—Lincoln, 200B Peter Kiewit Institute, 1110 South 67th Street, Omaha, Nebraska 68182-0681

Chair's Introduction—7:40

Invited Papers

7:45

1aAA1. Assessment of impact of acoustic and nonacoustic parameters on performance and well-being. Volker Mellert, Reinhard Weber (Inst. for Phys., Acoust., Oldenburg Univ., D-2611, Oldenburg, Germany), and Christian Nocke (Akustikbuero Oldenburg, Germany)

It is of interest to estimate the influence of the environment in a specific work place area on the performance and well-being of people. Investigations have been carried out for the cabin environment of an airplane and for class rooms. Acoustics is only one issue of a variety of environmental factors, therefore the combined impact of temperature, humidity, air quality, lighting, vibration, etc. on human perception is the subject of psychophysical research. Methods for the objective assessment of subjective impressions have been developed for applications in acoustics for a long time, e.g., for concert hall acoustics, noise evaluation, and sound design. The methodology relies on questionnaires, measurement of acoustic parameters, ear-related signal processing and analysis, and on correlation of the physical input with subjective output. Methodology and results are presented from measurements of noise and vibration, temperature and humidity in aircraft simulators, and of reverberation, coloring, and lighting in a primary school, and of the environmental perception. [The work includes research with M. Klatte, A. Schick from the Psychology Department of Oldenburg University, and M. Meis from Hoerzentrum Oldenburg GmbH and with the European Project HEACE (for partners see www.heace.org).]

8:05

1aAA2. A review of the combined effects of thermal and noise conditions on human performance. Richard A. Moscoso, Lily M. Wang, and Amy Musser (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, rmoscosobullon@mail.unomaha.edu)

Human perception and annoyance due to background noise has been the subject of much research. A great deal of work has also been done to identify conditions that produce an acceptable thermal environment for building occupants. The experience of occupants in indoor environments, however, is much more complex than can be represented by thermal comfort or the acoustic environment in isolation. Occupants normally experience a mix of thermal, auditory, visual, and olfactory stimuli that combines to form an impression of the environment. This paper is specifically interested in how building occupants trade off between acoustic and thermal comfort.

Heating, ventilation, and air-conditioning systems in buildings are often adjusted by building users to arrive at a more comfortable temperature, but this change may also produce more noise. Previous studies on the interaction effects between temperature and noise on human performance are reviewed in this presentation, followed by a discussion of the authors' current work in this area.

8:25

1aAA3. Ventilation noise and its effects on annoyance and performance. Ulf Landstrom (Natl. Inst. for Working Life North, Box 7654, SE-907 13 Umea, Sweden, ulf.landstrom@arbetslivsinstitutet.se)

In almost every room environment, ventilation acts as a more or less prominent part of the noise exposure. The contribution to the overall sound environment is a question not only of the way in which the ventilation system itself functions, but also a question of the prominence of other contemporary sound sources such as speech, equipment, machines, and external noises. Hazardous effects due to ventilation noise are most prominent in offices, hospitals, control rooms, classrooms, conference rooms, and other types of silent areas. The effects evoked by ventilation noise have also been found to be related to the type of activity being conducted. Annoyance and performance thus not only seemed to be linked to the physical character of exposure, i.e., noise level, frequency characteristics, and length of exposure, but also mental and manual activity, complexity, and monotony of the work. The effects can be described in terms of annoyance, discomfort, and fatigue, with consequences on performance and increased mental load. The silent areas where ventilation noise may be most frequently experienced are often synonymous with areas and activities most sensitive to the exposure.

8:45

1aAA4. The relevance of low-frequency sound properties for performance and pleasantness. Kerstin Persson Waye and Johanna Bengtsson (Dept. of Environ. Medicine, The Sahlgrenska Acad. of Gothenburg Univ., Box 414, 405 30 Gothenburg, Sweden, kerstin.persson-waye@envmed.gu.se)

The sound environment in the workplace has been found to influence performance, stress, mood, and well-being after work. However few studies can provide dose-response relationships and little is known of the importance of sound-quality aspects for adverse effects on critical tasks or task requirements. We have, during the last 8 years, been engaged in studies investigating the critical performance effects due to the presence of low frequencies (20–200 Hz) in sounds. The main hypotheses on critical effects derived from studies in the general environment were that low-frequency noise induced great annoyance, concentration difficulties, and was difficult to filter out or habituate to. On the other hand, results from truck drivers indicated that low-frequency sounds may lead to reduced alertness and increased sleepiness. In total, three studies were designed with regard to these hypotheses, all of them with the intention to be applicable to office and control room environment, using equivalent A-weighted sound-pressure levels of 40 and 45 dB. The fourth study investigated the importance of sound properties in low-frequency sounds for the perception of pleasantness. The results will be presented and discussed in relation to noise assessment aspects. [Work supported by Swedish Council for Working Life and Social Research.]

9:05

1aAA5. Relating human productivity and annoyance to indoor noise criteria systems. Erica E. Bowden and Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska−Lincoln, 245 PKI, 1110 S. 67th St., Omaha, NE 68182, ebowden@mail.unomaha.edu)

The goal of this research is to determine a noise criteria system which best relates the effects of background noise to human productivity and annoyance. A number of indoor noise criteria systems are currently used to rate the background noise in built environments, including noise criteria (NC), balanced noise criteria (NCB), room criteria (RC), room criteria Mark II (RC-Mark II), and others. Many questions still remain about the accuracy of these predictors in assessing human response to background noise under the variety of ambient noise situations encountered. To support the use of any individual criterion, subjective testing was performed under a range of background noise situations and statistically related to the various noise criteria predictors listed above. Subjects completed an annoyance survey and performed typing and proofreading tasks in an acoustically controlled environment under 12 simulated background noise settings. These settings varied across three sound levels and four spectral qualities. Subjective testing methodology and results are presented. [Work supported by INCE and ASHRAE.]

9:25

1aAA6. An investigation of the effects of interference speech on short-term memory for verbally presented prose. Dana M. Lodico (Illingworth & Rodkin, Inc., 505 Petaluma Blvd. S., Petaluma, CA 94952), Rendell R. Torres, Yasushi Shimizu, and Claudia Hunter (Rensselaer Polytech. Inst., Troy, NY 12180-3590)

This study investigates the effects of interference speech and the built acoustical environment on human performance, and the possibility of designing spaces to architecturally meet the acoustical goals of office and classroom environments. The effects of room size, geometry, and acoustical parameters on human performance are studied through human subject testing. Three experiments are used to investigate the effects of distracting background speech on short-term memory for verbally presented prose under constrained laboratory conditions. Short-term memory performance is rated within four different acoustical spaces and five background noise levels, as well as a quiet condition. The presentation will cover research methods, results, and possibilities for furthering this research. [Work supported by the Program in Architectural Acoustics, School of Architecture, Rensselaer Polytechnic Institute.]

9:45

1aAA7. Environmental impact on workplace performance. Derrick P. Knight, Steven J. Orfield, and Thomas J. Smith (Orfield Labs., 2709 E. 25th St., Minneapolis, MN 55406, derrick@orfieldlabs.com)

Since 1998, Orfield Laboratories has led research efforts into the effects of the office environment on worker performance through the Open Plan Working Group. This collaboration of researchers, designers, and facility managers works with OPWG in gathering environmental data (including noise, lighting, and thermal measurements) and subjective occupant ranking data. The OPWG then employs statistical analysis to correlate the environmental measurements, occupant performance, and occupant preference measurements. Through efficient measurement stations and Web-based surveys, Orfield Laboratories has helped many companies to begin to view their employees as their greatest operational profit center. This presentation will discuss the effect that noise as well as other environmental factors has on work performance.

10:05–10:15 Break

Contributed Papers

10:15

1aAA8. Impact of classroom noise on reading and vocabulary skills in elementary school-aged children. Prudence Allen, Nashlea Brogan, and Chris Allan (Natl. Ctr. for Audiol., Univ. of Western Ontario, London ON, Canada)

Classroom noise levels often exceed recommendations and, in large scale retrospective studies, it has been suggested that higher noise levels often correlate significantly with poorer academic performance [e.g., Shield., *et al.* (2002)]. However, experimental data on the performance of individual children are limited. This study therefore examined the effect of noise on the performance of children in grades 3–4 and 7–8 on standardized tests of oral reading, silent reading, and vocabulary (the Gray Oral Reading Test, the Gray Silent Reading Test, and the Peabody Picture Vocabulary Test). Each child completed parallel forms of a test in quiet and in classroom noise presented at 60 dB SPL. Required speech was presented at +10 S/N. Results from grouped data showed significantly reduced performance in noise only on the silent reading task and only for the older group of children. However, across tasks, when the effects of noise were evaluated as a function of childrens quiet performance levels, the noise effect was shown to be significant for the children performing at above average levels in quiet. These findings suggest that the effect of classroom noise may vary significantly across tasks and children. [Work supported by CLLRNet.]

10:30

1aAA9. Classroom acoustics in public schools: A case study. Carmen P. Loro and Paulo T. Zannin (Lab Acustica Ambiental, Universidade Federal do Parana, Dept. Engenharia Mecanica, Centro Politecnico, Curitiba, Brazil)

The acoustic quality of a standard classroom (Standard 23) of the public school system in the city of Curitiba has been evaluated. This standard has a central circulation aisle with two classrooms in each side. Each room has windows to the outside and to the internal aisle. Additionally, the aisle has a 6-m-high zenithal skylight, together composing the building's main lighting and ventilation system. But, Standard 23 lacks acoustic quality of the classrooms. In order to assay this, measurements have been performed under several conditions, using the Building Acoustics System of Bruel & Kjaer. The measured reverberation time (RT) of the four classrooms for a frequency of 500 Hz was: 1.65 s (empty classroom), 1.15 s (20 students in the room), and 0.76 s (40 students). According to WHO recommendations, the ideal RT in classrooms should be around 0.6 s. DIN 18041 establishes an RT between 0.8 and 1.0 s, to allow for adequate intelligibility. Background noise in an empty room was 63.3 dB (A), above the limit established by the Brazilian standard of acoustic comfort: 40 dB (A). The reaction of students and teachers has indicated that the main source of acoustic discomfort is the noise generated by the neighboring classrooms.

10:45

1aAA10. Marking emergency exits and evacuation routes with sound beacons utilizing the precedence effect. Sander J. van Wijngaarden, Adelbert W. Bronkhorst, and Louis C. Boer (TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands)

Sound beacons can be extremely useful during emergency evacuations, especially when vision is obscured by smoke. When exits are marked with suitable sound sources, people can find these using only their capacity for directional hearing. Unfortunately, unless very explicit instructions were given, sound beacons currently commercially available (based on modulated noise) led to disappointing results during an evacuation experiment in a traffic tunnel. Only 19% out of 65 subjects were able to find an exit by ear. A signal designed to be more self-explanatory and less hostile-sounding (alternating chime signal and spoken message "exit here") increased the success rate to 86%. In a more complex environment—a mock-up of a ship's interior—routes to the exit were marked using multiple beacons. By applying carefully designed time delays between successive beacons, the direction of the route was marked, utilizing the precedence effect. Out of 34 subjects, 71% correctly followed the evacuation route by ear (compared to 24% for a noise signal as used in commercially available beacons). Even when subjects were forced to make a worst-case left–right decision at a T-junction, between two beacons differing only in arrival of the first wave front, 77% made the right decision.

11:00

1aAA11. Relationships between objective acoustic indices and acoustic comfort evaluation in nonacoustic spaces. Jian Kang (School of Architecture, Sheffield Univ., Western Bank, Sheffield S10 2TN, UK, j.kang@sheffield.ac.uk)

Much attention has been paid to acoustic spaces such as concert halls and recording studios, whereas research on nonacoustic buildings/spaces has been rather limited, especially from the viewpoint of acoustic comfort. In this research a series of case studies has been carried out on this topic, considering various spaces including shopping mall atrium spaces, library reading rooms, football stadia, swimming spaces, churches, dining spaces, as well as urban open public spaces. The studies focus on the relationships between objective acoustic indices such as sound pressure level and reverberation time and perceptions of acoustic comfort. The results show that the acoustic atmosphere is an important consideration in such spaces and the evaluation of acoustic comfort may vary considerably even if the objective acoustic indices are the same. It is suggested that current guidelines and technical regulations are insufficient in terms of acoustic design of these spaces, and the relationships established from the case studies between objective and subjective aspects would be useful for developing further design guidelines. [Work supported partly by the British Academy.]

11:15

1aAA12. Allowable floor-impact sound levels in apartment buildings. Jin Yong Jeon and Jeong Ho Jeong (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea, jyjeon@hanyang.ac.kr)

The purpose of this study is to review the sound-isolation efficiencies of structural materials in multistory residential buildings, in which floor-impact sound was considered as the major source of residents' complaints. Floor, walls, and ceiling were selectively treated, and the floor-impact sound was measured using standard noise sources such as the impact ball, the bang machine, and the tapping machine. The noise from the impactors was analyzed and the relationship between the sound levels and the subjective responses was investigated. Results showed that the overall sound level of the impact ball is slightly higher than that of a bang machine, although the impact ball has a lower impact force. It was also found that the noise from the impact ball is similar to the noise of children running and jumping. In addition, when the noise from the impact ball was evaluated in both laboratory and *in situ* conditions, the allowable noise level was found to be 54 dB (Li, $F_{\max}$, AW).

11:30

1aAA13. Human annoyance and reactions to hotel room specific noises. Ian L. Everhard (Acoust. Dept., Columbia College Chicago, 33 E. Congress, Chicago, IL 60605)

A new formula is presented where multiple annoyance sources and transmission loss values of any partition are combined to produce a new single number rating of annoyance. The explanation of the formula is based on theoretical psychoacoustics and survey testing used to create variables used to weight the results. An imaginary hotel room is processed through the new formula and is rated based on theoretical survey results that would be taken by guests of the hotel. The new single number rating compares the multiple sources of annoyance to a single imaginary unbiased source where absolute level is the only factor in stimulating a linear rise in annoyance [Fidell *et al.*, J. Acoust. Soc. Am. **66**, 1427 (1979); D. M. Jones and D. E. Broadbent, "Human performance and noise," in *Handbook of Noise Control*, 3rd ed., edited by C. M. Harris (ASA, New York, 1998), Chap. 24; J. P. Conroy and J. S. Roland, "STC Field Testing and Results," in *Sound and Vibration Magazine*, Acoustical Publications, pp. 10–15 (July 2003)].

11:45

1aAA14. Investigation of comfort related aspects of noise in an aircraft cabin simulator. Reinhard Weber, Ingo Baumann, Nils Freese, and Volker Mellert (Inst. of Phys., Faculty V, Oldenburg Univ., D-26111 Oldenburg, Germany)

In the frame of the multinational European project HEACE Health effects of aircraft cabin environment [www.heace.org] experiments have been carried out to investigate the effects of different environmental factors in an aircraft on performance, comfort and health of flight and cabin crew. Tests were run in aircraft cabin simulators where temperature, humidity and sound could be adjusted in a controlled manner because only limited possibility exists of systematically changing these factors in-flight. In a multi-factorial $3\times3\times3$ design these tests simulated real flights with real cabin crew that was hired for the test and passenger. The research on passengers responses was done in cooperation with the European FACE Technology Platform (FACE Friendly aircraft cabin environment). This paper focuses on the effects of noise on the comfort on the cabin crew. It presents unexpected order effects of noise assessments and reports on the dependency of the ratings of noise and of other environmental factors on the assessed comfort. [The investigation is granted by the EU-Commission under HEACE G4RC-CT-2001-00611.]

MONDAY MORNING, 24 MAY 2004 VERSAILLES BALLROOM, 8:15 TO 11:55 A.M.

Session 1aAB

Animal Bioacoustics: Natural Acoustic Behavior of Animals: Session in Memory of Donald R. Griffin I

James A. Simmons, Chair
Department of Neuroscience, Brown University, Box 1953, Providence, Rhode Island 02912

Chair's Introduction—8:15

Invited Papers

8:20

1aAB1. Some recollections of D. R. Griffin as a young man. Robert Galambos (Dept. of Neurosci., UCSD, 8826 La Jolla Scenic Dr., La Jolla, CA 92037)

In 1939 Don Griffin invited me to join him in his earliest bat echolocation experiments. I will tell a few stories about what we two graduate students did together, and show the sound movie in which, for the first time, we recorded their cries as they flew and avoided obstacles.

8:40

1aAB2. Variability of feeding buzzes in little brown bats (*Myotis lucifigus*). Donald R. Griffin and Gregory J. Auger (Harvard Univ., Cambridge, MA)

When *Myotis lucifigus* are hunting actively in the early evening, search phase echolocation signals are easily detected by heterodyne bat detectors. Feeding or terminal buzzes are sometimes also detected, especially if the bat detector is tuned to 35 kHz. On other evenings when all conditions appeared comparable we detected no approach phase or buzz. Feeding bats and prey insects were observed and recorded with video and heterodyne and time expansion bat detectors during their early evening hunting at a small pond. Video was obtained with a Canon XL1 camcorder fitted with an ITT Pocketscope model 6010B light intensifier and near infra-red

light from the side. Audio was captured with a Pettersson D-980 bat detector using both the heterodyne and the time expansion outputs. Signals were recorded on the left and right audio tracks of the camcorder. In addition to recording general insect feeding, bats were offered at times a small tethered fly-fishing lure with the hook removed. Microphone to lure and insect distance was 0.5 to 2.0 m. Observations have shown variability in the length, presence, and loudness of search, approach, and terminal phases. Examples of video records of insect catching will be shown.

9:00

1aAB3. Bat echolocation calls: Orientation to communication. M. Brock Fenton (Dept. of Biol., Univ. of Western Ontario, London, ON N6A 5B7, Canada, bfenton@uwo.ca)

Bats hunting flying insects adjust the design of their echolocation calls according to the situation in which they forage and stage in an attack. Changes in call design across attack sequences alert other bats within earshot to the presence of prey, demonstrating a continuum in roles for biosonar signals between orientation and communication. Many aerial-feeding bats change the design of their echolocation calls in the presence of echolocating conspecifics. Bats may change frequency parameters, durations, and/or intensities of their calls. While a variety of free-tailed bats (*Molossidae Otomops martiensseni, Tadarida teniotis, Molossus molossus*) consistently change their echolocation calls when more than one bat is flying in an area, at least one sheath-tailed bat (*Emballonuridae Taphozous perforatus*) does not. Changes in echolocation calls may maximize jamming avoidance and/or enhance the communicative function of the calls. The data for molossids support the hypothesis that when hunting some species fly in formation. Here, variation in individual call design could provide positional information and reduce the chances of mid-air collisions.

9:20

1aAB4. Vocal communication of wild parrots. Jack Bradbury (Cornell Lab. of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850, jwb25@cornell.edu)

Field studies of four sympatric parrot species in Costa Rica are revealing several possible functions for the well-known ability of parrots to mimic new sounds throughout life. Despite earlier suggestions that this might facilitate exchanges of environmental information, all data so far suggest that vocal mimicry in the wild is associated with mediation of the fission/fusion of groups of parrots and/or of conflicts between mated pairs. Recent results using array recording and interactive playback will be summarized, and several technical problems created by the mechanisms of parrot vocal signal production discussed. [Research supported by NSF Grant IBN-022927 and by continued encouragement and logistics provided by the staff of the Area Conservacion Guanacaste (Costa Rica).]

9:40

1aAB5. Performance constraints and the production of birdsong. Roderick A. Suthers (Medical Sci., Indiana Univ., Bloomington, IN 47405), Eric Vallet (Univ. of Paris 10, Nanterre 92001, France), and Sue Anne Zollinger (Indiana Univ., Bloomington, IN 47405)

The role of physical and physiological constraints in determining the performance limits on the tempo and frequency bandwidth of birdsong was investigated. One series of experiments examined the mechanism by which a vocal mimic, the northern mockingbird (*Mimus polygottos*), copied the songs of other species with which it was tutored as a juvenile. Other experiments analyzed the motor basis of special canary (*Serinus canaria*) syllables eliciting sexual responses from females. In each case, the mechanism of vocalization was determined by measuring the respiratory dynamics and sound produced on each side of the songbirds duplex vocal organ, the syrinx. When mockingbirds copied the songs of other species the accuracy of their copy depended on the accuracy with which they reproduced the motor pattern used by the tutor species. Motor difficulty of various acoustic features was assessed by the accuracy of its copy. The high repetition rate, broadband canary syllables preferred by females required especially demanding bilateral motor skills. The results indicate that constraints on the rate of respiratory ventilation and bilateral syringeal coordination can set an upper limit on syllable repetition rate and frequency bandwidth. [Work supported by NIH and NSF.]

10:00–10:15 Break

10:15

1aAB6. Echolocation in wild toothed whales. Peter L. Tyack, Mark Johnson, Peter Teglberg Madsen (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, ptyack@whoi.edu), and Walter M. X. Zimmer (NATO Undersea Res. Ctr., 19138 La Spezia, Italy)

Don Griffin showed more than 50 years ago that bats echolocate for orientation and to capture prey. Experiments also demonstrated that captive dolphins can echolocate; more recent work parallels Griffin's work with bats in the wild. Digital acoustic recording tags were attached to sperm and beaked whales, *Ziphius cavirostris* and *Mesoplodon densirostris*, to record outgoing clicks and incoming echoes. The sperm whale data show echoes from the sea surface and seafloor, which are probably used for orientation and obstacle avoidance. When diving, sperm whales adjust their interclick interval as they change their pitch angle, consistent with the hypothesis that they are echolocating on a horizontal layer at the depth at which they will feed. This suggests that they may be listening for volume reverberation to select a prey patch. The beam pattern of sperm whales includes a narrow, forward-directed high-frequency beam probably used for prey detection, and a broader, backward-directed lower-frequency beam probably used for orientation. Beaked whales produce directional clicks with peak frequencies in the 25–40-kHz region. Echoes from individual prey items have been detected from clicks of beaked whales. This opens a new window into the study of how animals use echolocation to forage in the wild.

10:35

1aAB7. Array measurement of echolocation signals on the melon of harbor porpoises (*Phocoena phocoena*). Whitlow Au, Kelly Benoit-Bird (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, wau@hawaii.edu), Ronald Kastelein (Sea Mammal Res. Co., Harderwijk, The Netherlands), and Ted Cranford (San Diego State Univ., San Diego, CA)

The melon of odontocetes has been hypothesized to be a focusing body that channels echolocation signals produced within the nasal region of the animal's head into the water. The acoustic field of two echolocating harbor porpoises (*Phocoena phocoena*) was measured at the melon's surface with an array of four broadband hydrophones embedded in suction cups. The clicks detected by each hydrophone were simultaneously digitized at a sampling rate of 500 kHz and stored on a PC. Digital still photographs of the array were taken before each echolocation trial to measure the hydrophone's position. The shape and dimensions of the melon were measured with a flexible shape-retaining measuring device. The axis of the echolocation beam was found to be approximately 5.6–6.1 cm from the edge of the animal's upper lip along the midline of the melon, which coincides with the axis of the low-density lipid core of the melon. Click amplitudes dropped off rapidly (12–14 dB) from the maximum at hydrophones 3.5 cm apart, providing support for the melon-focusing hypothesis. Changes in the waveform on consecutive hydrophones suggest that the porpoises can manipulate either the shape of the melon or the output of the sources within 0.086 s.

10:55

1aAB8. Echolocation click rates and behavior of foraging Hawaiian spinner dolphins. Kelly J. Benoit-Bird and Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

Groups of spinner dolphins work together to actively aggregate small animals in the deep-scattering layer that serve as their prey. Detailed information on dolphin foraging behavior, obtained with a 200-kHz multibeam sonar (Simrad MS2000), made it possible to correlate echolocation and foraging. Fifty-six groups of spinner dolphins foraging at night within a midwater micronekton sound-scattering layer were observed with the sonar. During sonar surveys, the rates of whistles and echolocation clicks were measured using four hydrophones at 6-m depth intervals. Significant differences in click rates were found between depths and between the different stages of foraging. Groups of foraging dolphins ranged in size from 16 to 28 dolphins. Click rates were not significantly affected by the number of dolphins in a foraging group. Contrary to initial predictions, click rates were relatively low when sonar data indicated that pairs of dolphins were actively feeding. Highest echolocation rates occurred within the scattering layer, during transitions between foraging states. Whistles were only detected when dolphins were not in a foraging formation and when animals were surfacing. This suggests clicks may be used directly or indirectly to cue group movement during foraging.

11:15

1aAB9. Stuttering: A novel bullfrog vocalization. Andrea Simmons and Dianne Suggs (Dept. of Psych., Brown Univ., Providence, RI 02912, Andrea_Simmons@brown.edu)

The advertisement call of male bullfrogs (*Rana catesbeiana*) consists of a series of individual croaks, each of which contains multiple harmonics with a missing or attenuated fundamental frequency of approximately 100 Hz. The envelope of individual croaks has typically been represented in the literature as smooth and unmodulated. From an analysis of 5251 advertisement calls from 17 different choruses over two mating seasons, we show that males add an extra modulation (around 4 Hz) to the envelope of individual croaks, following specific rules. We term these extra modulations stutters. Neither single croak calls nor the first croak in multiple croak calls contains stutters. When stuttering begins, it does so with a croak containing a single stutter, and the number of stutters increases linearly (plus or minus 1 stutter, up to 4 stutters) with the number of croaks. This pattern is stable across individual males ($N = 10$). Playback experiments reveal that vocal responses to stuttered and nonstuttered calls vary with proximity to the stimulus. Close males respond with nonstuttered calls, while far males respond with stuttered calls. The data suggest that nonstuttered calls are used for aggressive or territorial purposes, while stuttered calls are used to attract females.

11:35

1aAB10. Impact of the chorus environment on temporal processing of advertisement calls by gray treefrogs. Joshua Schwartz, Kenneth Huth, and Jeffrey Lasker (Dept. of Biol., Pace Univ., 861 Bedford Rd., Pleasantville, NY 10570, jschwartz2@pace.edu)

Male gray treefrogs advertise for mates using calls that consist of a series of pulses. Pulse duration, interpulse interval, and pulse shape determine whether a call is recognized as a conspecific signal by females. Females use call rate and call pulse number to assess relative calling performance by males, and prefer males that display high calling efforts. However, within choruses call overlap among males and background noise can compromise the ability of females to detect and correctly interpret temporal information in calls. Phonotaxis tests using calls suffering from different patterns of overlap or with internal gaps were used to investigate specific consequences of interference and masking as well as mechanisms that might alleviate such problems. Our data indicate that females do not employ a process analogous to phonemic restoration to "fill in" missing call segments; however, if a sufficient percent of call elements fall within species-specific ranges, females may ignore call anomalies. Additional findings are generally consistent with those from a recent study on anuran auditory midbrain neurons that count and indicate that inappropriate pulse intervals can reset the pulse counting process. [Work supported by NSF and a Pace University Eugene M. Lang Research Fellowship.]

Session 1aBB

Biomedical Ultrasound/Bioresponse to Vibration and Signal Processing in Acoustics: High Frequency Imaging

Jeffrey A. Ketterling, Cochair
Riverside Research Institute, 156 William Street, New York, New York 10038-2609

Orlando Aristiz ábal, Cochair
Skirball Institute of Bimolecular Medicine, 540 First Avenue, New York, New York 10016

Chair's Introduction—7:55

Invited Papers

8:00

1aBB1. Current and future innovations in high-frequency ultrasonic transducers and arrays. K. Kirk Shung (Dept. of Biomed. Eng., Univ. of Southern California, Los Angeles, CA 90089, kkshung@usc.edu)

High-frequency ultrasonic imaging is considered the next frontier in ultrasonic imaging. Commercial high-frequency scanners often termed ultrasonic backscatter microscope (UBM) all use scanned single element transducers at frequencies higher than 30 MHz with a frame rate lower than 30 frames/s. The engineering of single element transducers and linear arrays at these frequencies has been problematic and may be addressed from the end of developing better piezoelectric materials and from the end of developing novel fabrication methodologies. The further improvement of UBM can be benefited greatly by the availability of superior single element transducers of wider bandwidth and higher sensitivity. Various novel piezoelectric materials, including 2-2 composites, fine grain materials, high or low dielectric constant materials, and fabrication methods that utilize conventional dice and fill, thick film, and MEMS have been explored to fulfill this need. Prototype linear arrays higher than 30 MHz have been developed and tested and preliminary images obtained. These recent developments will be discussed in this talk. [Work supported by NIH.]

8:20

1aBB2. Optoacoustics for high-frequency ultrasonic imaging and manipulation. Matthew O'Donnell and Takashi Buma (Biomed. Eng. Dept. and Appl. Phys. Prog., Biomed. Ultrason. Lab., Univ. of Michigan, Ann Arbor, MI 48109-2099)

Pulsed lasers can generate ultrasound through thermoelastic expansion of a thin optical absorber. By carefully designing the optical absorbing structure, efficient transduction is possible for a number of biomedical applications including high-frequency imaging, microfluidics, and sensing. The major key for efficient optoacoustic transduction in biomedical applications is to engineer a nearly perfect optical absorber possessing a large coefficient of thermal expansion with acoustic properties well matched to a water medium. We have obtained an optoacoustic efficiency increase of over 20 dB compared to conventional approaches using a thin, optically absorbing layer consisting of polydimethylsiloxane (PDMS) and carbon black spin coated onto a clear PDMS substrate. This structure has been extensively analyzed both experimentally and analytically and seems to provide opportunities for a wide range of optoacoustic devices. In this talk we show how PDMS-based optoacoustic transduction can be used for high-frequency imaging using longitudinal waves and acoustic tweezing using Lamb waves. The basic mechanism of optoacoustic transduction will be described, and specific devices will be presented.

8:40

1aBB3. Ultrasound biomicroscopy in mouse cardiovascular development. Daniel H. Turnbull (Skirball Inst. of Biomolecular Medicine and Depts. of Radiol. & Pathol., New York Univ. School of Medicine, New York, NY 10016, turnbull@saturn.med.nyu.edu)

The mouse is the preferred animal model for studying mammalian cardiovascular development and many human congenital heart diseases. Ultrasound biomicroscopy (UBM), utilizing high-frequency (40–50-MHz) ultrasound, is uniquely capable of providing *in vivo*, real-time microimaging and Doppler blood velocity measurements in mouse embryos and neonates. UBM analyses of normal and abnormal mouse cardiovascular function will be described to illustrate the power of this microimaging approach. In particular, real-time UBM images have been used to analyze dimensional changes in the mouse heart from embryonic to neonatal stages. UBM-Doppler has been used recently to examine the precise timing of onset of a functional circulation in early-stage mouse embryos, from the first detectable cardiac contractions. In other experiments, blood velocity waveforms have been analyzed to characterize the functional phenotype of mutant mouse embryos having defects in cardiac valve formation. Finally, UBM has been developed for

real-time, *in utero* image-guided injection of mouse embryos, enabling cell transplantation and genetic gain-of-function experiments with transfected cells and retroviruses. In summary, UBM provides a unique and powerful approach for *in vivo* analysis and image-guided manipulation in normal and genetically engineered mice, over a wide range of embryonic to neonatal developmental stages.

9:00

1aBB4. Clinical applications of very high frequency ultrasound in ophthalmology. Ronald H. Silverman, D. Jackson Coleman, Dan Z. Reinstein (Dept. of Ophthalmology, Weill Medical College of Cornell Univ., 1300 York Ave., New York, NY 10021), and Frederic L. Lizzi (Riverside Res. Inst., New York, NY 10038)

The eye is ideally suited for diagnostic imaging with very high frequency (>35 MHz) ultrasound (VHFU) because of its peripheral location and cystic structure. VHFU allows high resolution visualization of pathologies affecting the anterior segment of the eye, including tumors, cysts, foreign bodies, and corneal pathologies. We developed a series of prototype instruments suitable for ophthalmic studies using both polymer and lithium niobate transducers, with digitization of radiofrequency echo data at up to 500 MHz. While initially using linear scan geometries, we subsequently developed an arc-shaped scan matched to the curvature of the 0.5-mm-thick cornea to circumvent the effect of specular deflection of the ultrasound beam produced by the corneas curved surface. This technique allowed us to obtain data across the entire cornea and determination of the thickness of each corneal layer, including the epithelium (approximately 50 microns in thickness) and the surgically induced interface produced in LASIK, the most common form of refractive surgery. By scanning in a series of meridians, and applying optimized signal processing strategies (deconvolution, analytic signal envelope determination), corneal pachymetric maps representing the local thickness of each layer can be generated and aid in diagnosis of surgically induced defects or refractive abnormalities.

9:20

1aBB5. Scattering and statistical models for very-high-frequency ultrasonic monitoring of tumor therapy. Frederic L. Lizzi (Riverside Res. Inst., 156 William St., New York, NY 10038)

Our laboratories examined spectral changes in ultrasonic backscatter from tumors undergoing various forms of treatment, with the goal of developing noninvasive treatment monitoring. Human tumor explants in athymic mice were treated *in vivo* with ultrasonic hyperthermia, high-intensity focused ultrasound, and chemical agents. Pre- and posttreatment spectral examinations were conducted using very-high frequency ultrasound (40-MHz center frequencies). Physical scatterer properties were estimated from measured spectral data. Spatially averaged spectra showed that successful treatment progressively increased acoustic concentration by about 3 dB. It also produced average changes of 2 μm in scatter sizes of cell-sized (10 μm) structures; these perturbations may be associated with apoptosis, vacuole formation, and frank cellular disruption. These findings are consistent with clinical observations (40 MHz) following radiotherapy of ocular tumors. Theoretical scattering and statistical models were applied to specify transducer and processing schemes that will permit these changes to be mapped in high resolution images, rather than evaluated as spatially averaged values. This requires improved estimator precision for spectral assays of small scatterers near the Rayleigh size limit. Results showed that this goal can be achieved with 50-MHz annular arrays combined with adaptation of spectral calibration and processing procedures; these are now being implemented for clinical application.

Contributed Papers

9:40

1aBB6. Imaging the shear flow of complex fluids using high-frequency ultrasound. Sébastien Manneville, Lydiane Bécu, Annie Colin (Ctr. de Recherche Paul Pascal—CNRS, Ave. Schweitzer, F-33600 Pessac, France), Mickaël Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, 75005 Paris, France)

Complex fluids show fascinating properties under flow due to the existence of a *mesoscopic* scale located between the microscopic and the macroscopic scales. For instance, uncoiling of polymer molecules under flow leads to well-documented shear-thinning behaviors. More surprisingly, *inhomogeneous* flows of visco-elastic fluids are observed in simple shear experiments even at very low Reynolds numbers. Indeed, due to strong coupling between the flow and the fluid microstructure, shear may induce new structural organizations that coexist in the flow leading to shear localization or to shear bands. Whereas classical rheology only yields *global* data such as the viscosity averaged over the whole sample, the present work is devoted to *local* ultrasonic velocimetry in sheared complex fluids. A 1D high-frequency (36 MHz) speckle tracking technique is presented that allows the spatio-temporal study of various inhomogeneous unsteady flows. Moreover, preliminary results using a 2D ultrafast flow imaging technique at 12 MHz show that images of the flow can be obtained with an unprecedented temporal resolution.

9:55

1aBB7. Breast tissue characterization with high-frequency scanning acoustic microscopy. R. E. Kumon, I. Bruno, B. Heartwell, and E. Maeva (Dept. of Phys., Univ. of Windsor, 401 Sunset Ave., Windsor, ON N9B 3P4, Canada, kumon@uwindsor.ca)

We have performed imaging of breast tissue using scanning acoustic microscopy (SAM) in the range of 25–50 MHz with the goal of accurately and rapidly determining the structure and composition throughout the volume of the samples. In contrast to traditional histological slides, SAM images can be obtained without special preparation, sometimes even without sectioning, but with sufficiently high spatial resolution to give information comparable to surface optical images. As a result, the use of high-frequency SAM at the time of breast lumpectomy to identify disease-free margins has the potential to reduce reoperative rates, patient anxiety, and local recurrence. However, only limited work has been performed to characterize breast tissue in the frequency range above clinical ultrasound devices. The samples are 4-cm^2-thick sections (2–3 mm) taken from mastectomies and preserved in formalin. They are placed between two plates and immersed in water during imaging. Attenuation images are acquired by focusing the acoustic beam at the top and bottom of the samples, although better results were obtained for bottom focusing. For purposes of comparison and identification of histological features, acoustical images will be presented along with optical images obtained from the same samples. [Work supported by CIHR.]

10:10–10:30 Break

10:30

1aBB8. *In vitro* measurements of acoustic properties using a high-frequency scanning acoustic microscope (SAM). Caleb H. Farny, T. L. Szabo, Ronald A. Roy, and Robin O. Cleveland (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

A scanning acoustic microscope (SAM) is described that can measure attenuation, sound speed, impedance, and backscatter coefficient of a sample. The SAM consists of a spherically focused transducer (10–40 MHz) which operates in pulse-echo mode and is scanned in a 2D raster pattern over a sample. A plane wave analysis is presented which allows the impedance, attenuation, and phase velocity of a sample to be recovered from the front and back echoes. A causality model is used to validate the attenuation, and phase velocity measurements. The analysis of Chen *et al.* [UFFC, **445**, 515] is used to obtain the backscatter coefficient based on the echoes generated by subwavelength scatterers within the sample. The results for the impedance, attenuation and phase velocity were validated for high- and low-density polyethylene against published results. The SAM was used to measure the impedance, attenuation, phase velocity, and backscatter coefficient for the medial and adventitial layers of an *in vitro* human femoral artery. The theory was extended to account for propagation through multiple layers. The SAM measurements of acoustic properties were in good agreement with previously published results. [Work supported by the NSF through the Center for Subsurface Sensing and Imaging Systems.]

10:45

1aBB9. Robust diffraction correction method for high-frequency ultrasonic tissue characterization. Balasundar Raju (Philips Res., 345 Scarborough Rd., Briarcliff Manor, NY 10510)

The computation of quantitative ultrasonic parameters such as the attenuation or backscatter coefficient requires compensation for diffraction effects. In this work a simple and accurate diffraction correction method for skin characterization requiring only a single focal zone is developed. The advantage of this method is that the transducer need not be mechanically repositioned to collect data from several focal zones, thereby reducing the time of imaging and preventing motion artifacts. Data were first collected under controlled conditions from skin of volunteers using a high-frequency system (center frequency=33 MHz, BW=28 MHz) at 19 focal zones through axial translation. Using these data, mean backscatter power spectra were computed as a function of the distance between the transducer and the tissue, which then served as empirical diffraction correction curves for subsequent data. The method was demonstrated on patients patch-tested for contact dermatitis. The computed attenuation coefficient slope was significantly ($p<0.05$) lower at the affected site (0.13 ± 0.02 dB/mm/MHz) compared to nearby normal skin (0.2 ± 0.05 dB/mm/MHz). The mean backscatter level was also significantly lower at the affected site (6.7 ± 2.1 in arbitrary units) compared to normal skin (11.3 ± 3.2). These results show diffraction corrected ultrasonic parameters can differentiate normal from affected skin tissues.

11:00

1aBB10. Reciprocity calibration of hydrophones at various temperatures in the MHz frequency range. Cecille Labuda, Charles C. Church, and Jason Raymond (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, cpembert@olemiss.edu)

Calibration of an Imotec 300/24/62 bilaminar PVDF needle hydrophone and a Sonora Medical S4-155 membrane hydrophone over the frequency range 1–12 MHz at different temperatures (3–37 °C) was conducted using a reciprocity method. Calibration data received from hydrophone manufacturers are given for a range of frequencies but usually at only one temperature, typically room temperature. If the hydrophone must be used at a temperature different than the one at which it was calibrated, a new calibration must be performed. The reciprocity method employed in this study makes use of pairs of unfocused reciprocal transducers (send/receive) to provide redundant measurements and simplicity. The method was tested by calibrating a hydrophone for which the manufacturers calibration data were available. Our test data, with error bars of about 1.7 dB, fell within the range of the manufacturers data, which have error bars of about 1 dB. The results for the passive PVDF probe (Imotec) showed that sensitivity decreased by about one order of magnitude as the temperature decreased from 22 °C (i.e., room temperature) to 3 °C. A small increase is observed as temperature rises from 22 °C to 37 °C although the sensitivity remains about the same order of magnitude.

11:15

1aBB11. High frequency nonlinear imaging. Jens K. Poulsen (The Inst. of Exp. Clin. Res., Aarhus Univ., 8200 Aarhus N, Denmark)

Nonlinear imaging has already shown improved image resolution compared to fundamental imaging at lower frequencies (2–4 MHz). The required steps to obtain nonlinear images at high frequencies are presented here. The transmit frequency was from 20 to 60 MHz and on receive, pulse-inversion, the first, second, and the third harmonic were used for imaging. Experiments were conducted using a wire phantom and *in vitro* from a human femoral artery. PVDF transducer (6 mm diameter, 10 mm focal length, central frequency 42 MHz, −6 dB bandwidth 38 MHz) was used with a filtered amplifier (46.0 and 50.6 dBm output). The pulse-inversion experiment at 40 and 60 MHz had lateral resolution (−6 dB contours) 64 and 52 μm and axial 43 and 31 μm. The fundamental experiments at 20, 30, 40, and 60 MHz had lateral resolutions 104, 89, 74, 57 μm and axial 72, 86, 46, 35 μm. The second harmonic at 40 and 60 MHz had lateral resolution 62 and 47 μm and axial 52 and 38 μm. The third harmonic at 60 MHz had lateral resolution 48 μm and axial 42 μm. Nonlinear tissue images had improved resolution compared to fundamental imaging, though the improvement was small for the second harmonic experiment.

11:30

1aBB12. 40–50-MHz lithium niobate ($LiNbO_3$) transducers for pulsed Doppler measurements in mouse embryos. Orlando Aristizábal, Daniel H. Turnbull, Ruiping Ji (Skirball Inst., 540 First Ave., New York, NY 10016, oarist@saturn.med.nyu.edu), and Colin Phoon (New York Univ. School of Medicine, New York, NY 10016)

An ultrasound biomicroscopy (UBM) system operating between 40 and 50 MHz has been developed for imaging and detecting blood flow in normal and genetically modified mouse embryos, which has contributed to the understanding of the functional consequences of specific genetic defects affecting cardiovascular development. In this scanner, UBM images and Doppler blood velocity waveforms are acquired simultaneously with separate transducers, enabling independent optimization of either the imaging or Doppler transducer. Air backed, 1.5-mm-diameter, unfocused 44-MHz $LiNbO_3$ Doppler transducers designed for the 40- to 50-MHz pulsed Doppler ultrasound system have been developed. The fabrication and characterization of the electromechanical and ultrasonic beam properties of these transducers will be described. The two-way insertion loss at 44 MHz was measured to be −16 dB before tuning, a 60% improvement over the piezo-polymer transducers typically used for UBM-Doppler. Beam-plot data demonstrate the expected collimated near-field pattern to a depth of 13 mm, providing reasonable lateral resolution (1.0–1.5 mm) over a wide depth of field. The utility of these Doppler transducers for interrogating blood vessels such as the dorsal aorta in normal and mutant mouse embryos with defects in cardiac valve formation has been demonstrated.

11:45

1aBB13. Design and testing of an annular array for very-high-frequency imaging. Jeffrey A. Ketterling, Sarayu Ramachandran, Frederic L. Lizzi (Riverside Res. Inst., 156 William St., New York, NY 10038, ketterling@rrinyc.org), Orlando Aristizábal, and Daniel H. Turnbull (New York Univ. School of Medicine, New York, NY 10016)

Very-high-frequency ultrasound (VHFU) transducer technology is currently experiencing a great deal of interest. Traditionally, researchers have used single-element transducers which achieve exceptional lateral image resolution although at a very limited depth of field. A 5-ring focused

annular array, a transducer geometry that permits an increased depth of field via electronic focusing, has been constructed. The transducer is fabricated with a PVDF membrane and a copper-clad Kapton film with an annular array pattern. The PVDF is bonded to the Kapton film and pressed into a spherically curved shape. The back side of the transducer is then filled with epoxy. One side of the PVDF is metallized with gold, forming the ground plane of the transducer. The array elements are accessed electrically via copper traces formed on the Kapton film. The annular array consists of 5 equal-area rings with an outer diameter of 1 cm and a radius of curvature of 9 mm. A wire reflector target was used to test the imaging capability of the transducer by acquiring B-scan data for each transmit/receive pair. A synthetic aperture approach was then used to reconstruct the image and demonstrate the enhanced depth of field capabilities of the transducer.

Session 1aNS

Noise: Noise Effects and Hearing Protection

Elliott H. Berger, Chair
E-A-R/Aearo Company, 7911 Zionsville Road, Indianapolis, Indiana 46268-1657

Chair's Introduction—10:10

Contributed Papers

10:15

1aNS1. The development of noise-induced hearing loss in military trades. Sharon M. Abel and Stephanie Jewell (Defence Res. and Development Canada—Toronto, 1133 Sheppard Ave. W., Toronto, ON M3M 3B9, Canada)

An investigation is in progress to determine risk factors for the development of noise-induced hearing loss in Canadian Forces personnel. A total of 1057 individuals representing a wide range of military trades have contributed their current audiogram, first audiogram on record, and responses to a 56-item questionnaire. The protocol for the hearing test was standardized and conformed to current audiological practice. The items included in the questionnaire related to demographics, occupational and nonoccupational noise exposure history, training in and utilization of personal hearing protectors, and factors other than noise which might affect hearing (e.g., head injury, ear disease, exposure to solvents, and the use of medications). Analyses are underway to determine the average current hearing thresholds as a function of frequency and change relative to baseline values at recruitment for groups defined by trade, rated noise hazard, and years of service. Preliminary results suggest ways to improve the training of personnel with respect to the effects of both occupational and nonoccupational noise exposure and methods of implementing hearing conservation strategies. The role of head injury, history of ear disease, and the use of medications appear to be small. [Work supported by Veterans Affairs Canada.]

10:30

1aNS2. Estimating the precision error in hearing protector ratings. William J. Murphy, John R. Franks (Hearing Loss Prevention Section, NIOSH, 4676 Columbia Pkwy., MS C-27, Cincinnati, OH 45226, wjm4@cdc.gov), and Peter B. Shaw (NIOSH, Cincinnati, OH 45226)

The variance associated with hearing protector rating methods is derived for the Noise Reduction Rating, Noise Reduction Rating Subject Fit, and Single Number Rating methods. Real-ear attenuation at threshold (REAT) data have an error that is a function of stimulus frequency, protector fit, protector type, and the subject performance during occluded and unoccluded threshold estimation. In addition to the error associated with conducting the REAT tests, there is an additional error that results from applying the rating to a noise spectrum different from the spectrum used to derive the rating. For this paper, a distinction is made between the accuracy of the rating and the precision of the rating. The precision is an inherent property of the subject panel, protector type and fit, and the frequencies tested. The accuracy will be attributed to the application of the rating to an arbitrary noise. [Portions of work supported by the U.S. EPA IA 75090527.]

10:45

1aNS3. A new hearing protector rating: The Noise Reduction Statistic for use with A weighting (NRS_A). Elliott H. Berger (E-A-R/Aearo Co., 7911 Zionsville Rd., Indianapolis, IN 46268-1657, eberger@compuserve.com) and Dan Gauger (Bose Corp., Framingham, MA 01701-9168)

An important question to ask in regard to hearing protection devices (HPDs) is how much hearing protection they can provide. With respect to the law, at least, this question was answered in 1979 when the U. S. Environmental Protection Agency (EPA) promulgated a labeling regulation specifying a Noise Reduction Rating (NRR) measured in decibels (dB). In the intervening 25 years many concerns have arisen over this regulation. Currently the EPA is considering proposing a revised rule. This report examines the relevant issues in order to provide recommendations for new ratings and a new method of obtaining the test data. The conclusion is that a Noise Reduction Statistic for use with A weighting (NRS_A), an A–A' rating computed in a manner that considers both intersubject and interspectrum variation in protection, yields sufficient precision. Two such statistics ought to be specified on the primary package label—the smaller one to indicate the protection that is possible for most users to exceed, and a larger one such that the range between the two numbers conveys to the user the uncertainty in protection provided. Guidance on how to employ these numbers, and a suggestion for an additional, more precise, graphically oriented rating to be provided on a secondary label, are also included.

11:00

1aNS4. Real-world performance of hearing protection systems in Army aviation. William A. Ahroon and Martin B. Robinette (U.S. Army Aeromedical Res. Lab., P.O. Box 620577, Fort Rucker, AL 36362-0577, william.ahroon@us.army.mil)

Methods to measure the performance of hearing protective devices using naive subjects were developed to better estimate the hearing protection that can be reasonably achieved in operational environments. A ''pref-

erence'' for subject-fit methods over experimenter-fit or experimenter-supervised-fit methods has been growing among many hearing conservation professionals but has not received universal acceptance. The real-world performance of hearing protection used in Army aviation was measured. Army aviators who had completed 38 weeks of rotary-wing aircraft training were used in personal hearing protector evaluations using ANSI S12.6 real-ear attenuation at threshold procedures. Using their own flight helmets (group 1) or helmets and earplugs (group 2), the aviators were instructed to fit their helmet or helmets with earplugs as they do for flight operations (i.e., an informed-user fit). Following this evaluation, an experimenter-supervised fit evaluation was performed, which included helmet and earplug fit by a technician trained and experienced in helmet fit procedures. For both groups, the real-ear attenuation at threshold results for the informed-user and experimenter-supervised fitting procedures were virtually identical. Thus, the experimenter-supervised fit procedure (ANSI S12.6-1997 Method A) is appropriate for evaluating the performance of personal hearing protective headgear in the Army aviation environments.

11:15

1aNS5. Attenuation of high-level acoustic impulses by earplugs. Jan Zera and Rafal Mlynski (Central Inst. for Labour Protection, Natl. Res. Inst., ul. Czerniakowska 16, 00-701 Warsaw, Poland)

Attenuation of acoustic impulses by various models of foam, premolded, formable and semi-insert earplugs was measured. The impulses were generated by a loudspeaker system in the 115–130 range of levels, and by a blast of air expanding from a cylinder in the 150–170 level range. Transmission loss method was used to determine the difference between the peak level under the earplug and outside the earplug. The earplugs were inserted into in-house designed cylindrical or conical couplers attached to an artificial test fixture. The measurements showed that attenuation of high-level impulses depends on the type of coupler and the depth of earplug insertion. Attenuation measured using a conical coupler was consistently about 5 dB lower than for cylindrical coupler. Decreasing the depth of earplug insertion from 75% to 50% was especially critical at high level impulses and lowered attenuation by about 6 dB. For various models of foam and premolded earplugs, attenuation of high-level impulses varies by more than 20 dB. Particularly high attenuation was obtained for formable earplug made of modified wax. [Work supported by the State Committee for Scientific Research Grant No. III-7.03, III-6.07.]

11:30

1aNS6. Evidence for training effects of the acoustical environment on hearing. Gerald Fleischer and Reinhard Mueller (Auditory Res., Univ. of Giessen, Aulweg 123, 35392 Giessen, Germany)

A database of roughly 10 000 adult persons—from 18 years to 70 years—was analyzed with a procedure that accounts for normal aging of the ear. Auditory performance of good-hearing persons of various groups was determined: office personnel, construction workers, university students, airline pilots, dentists, orchestra musicians, fans of discotheques, avoiders of discotheques, Tibetian nomads, Chinese peasants living without technical noise, etc. Pure-tone auditory threshold—based on pulsed signals—was analyzed from 125 Hz to 10 kHz. The Leq of the long-term acoustic environment was estimated, using acoustic measurements. Results confirm the well-known fact that excessive noise levels damage the ear. However, at lower levels, sound can apparently improve the sense of hearing, as measured by the auditory threshold. Ranking the various groups according to their hearing performance reveals that the best-hearing groups are all living and working in a loud acoustic environment. In these groups aging of the ear is reduced. Groups living in an environment with very low sound levels do not hear well. Apparently the auditory system needs training, in order to develop its full potential, and to keep functioning well.

11:45

1aNS7. Vulnerability of the ear in men and women. Gerald Fleischer and Reinhard Mueller (Auditory Res., Univ. of Giessen, Aulweg 123, 35392 Giessen, Germany)

A special procedure was developed to separate ears with auditory damage from those without damage. It was applied to a database of roughly 10 000 persons, containing pure-tone audiograms and other information. Persons with medical problems of the ear were excluded. For analysis, groups of persons with similar ways of life were examined: college students, orchestra musicians, dentists, Tibetian nomads, Chinese city dwellers, etc. Looking specifically at ears that suffered noise-induced auditory damage, the damages are different in men and women. At frequencies below about 2 kHz men typically have less hearing loss than women. However, above 2 kHz, men have more hearing loss than women. This effect is quite strong and highly significant. It could be found in every group that is large enough. It is present in persons going regularly to discotheques, as well as in Tibetian nomads, who live without any technical noise, but use fire crackers now and then. This effect appears not related to the acoustic environment, but can be the result of differences in vulenerability between men and women. It can also be found in youngsters and, to some extent, in children.

Session 1aPAa

Physical Acoustics: Thermoacoustics, Bubbles and Films

David L. Gardner, Chair
Los Alamos National Laboratory, Condensed Matter and Thermal Physics Group, Los Alamos, New Mexico 87545

Contributed Papers

8:00

1aPAa1. DeltaEZ: A visual interpreter for DeltaE. Gordon P. Smith (Western Kentucky Univ., Bowling Green, KY 42101)

Thermoacoustic devices typically begin their development utilizing DeltaE, a numerical simulator developed by researchers at the Los Alamos National Laboratories. The program is highly popular amongst the thermoacoustics community for its ability to successfully describe the physical parameters and geometries that a thermoacoustic device requires to operate. However, the program is text-based, and can be somewhat counterintuitive to the novice user. Alternate, graphically based, programs have been developed for the thermoacoustic community, meeting with mixed reactions. This approach differs from earlier efforts in that no attempt is made to replace DeltaE itself. A front-end interface possessing the intuitive benefits of graphical programs has been created, which then runs DeltaE. Details of the beta version of the program will be presented in the hopes of gathering volunteers for further development.

8:15

1aPAa2. A small-scale, thermoacoustic-Stirling electric generator for deep-space applications. Scott Backhaus (Condensed Matter and Thermal Phys. Group, Los Alamos Natl. Lab., Los Alamos, NM 87545), Michael Petach, and Emanuel Tward (Northrop Grumman Space Technol., Redondo Beach, CA 90277)

Although thermoacoustic-Stirling hybrid engines (TASHE) have not been previously coupled to transducers to produce useful electric power, they have demonstrated high thermal-to-acoustic power conversion efficiencies. Electric generation is investigated by coupling a small TASHE to an electrodynamic linear alternator with an emphasis on satisfying NASA's need for a small, lightweight, efficient electric generator for deep-space missions. The combined goals of low mass and high efficiency require the TASHE to have the largest acoustic power output possible from a minimum enclosed volume, which imposes a relation between various impedances of the TASHE's lumped-element loop. The design of the TASHE and alternator used in this generator will be reviewed, performance data presented, and possible improvements discussed. [Work supported by NASA.]

8:30

1aPAa3. Characterization of regenerator materials for thermoacoustic refrigeration. Jin Liu and Steven Garrett (Grad. Prog. in Acoust., Penn State, P.O. Box 30, State College, PA 16804)

A bellows bounce test apparatus using a mechanically resonant gas-filled metal bellows, coupled to moving-magnet linear motor, contained within a pressure vessel, has been developed for the testing of porous regenerator materials that might be useful for thermoacoustic refrigeration. The apparatus provides a combination of operating pressures, gases, amplitudes, and frequencies that can cover a large range of Reynolds numbers and Lautrec numbers (ratio of the hydraulic radius to the thermal penetration depth). We will report initial results obtained with this apparatus using two ceramic Celcor samples (400 and 600 cells/in*in) over four orders-of-magnitude in Reynolds number. The experimental results were in good agreement with the linear acoustic approximation at low Reynolds numbers, but deviate at higher Reynolds numbers for the viscous component of the pressure drop. For low Lautrec numbers, the dimensionless pressure drop depends on both Lautrec and Reynolds numbers. For Lautrec numbers greater than one (Poiseuille flow regime), there is no frequency (Lautrec) dependence. Similitude analysis collapses all data on both viscous (in-phase) and inertial (quadrature) pressure drops suggesting that introduction of the Lautrec number (or some dimensionless variable related to penetration depth) was required and measurements on samples of more complicated geometries should provide reliable results.

8:45

1aPAa4. Thermoacoustic boundary layers near the liquid–vapor critical point. Keith A. Gillis, Iosif I. Shinder, and Michael R. Moldover (Process Measurements Div., NIST, Gaithersburg, MD 20899-8360, keith.gillis@nist.gov)

The sound attenuation in resonators filled with xenon at its critical density ρ_c was calculated and measured as a function of the reduced temperature $\tau \equiv (T - T_c)/T_c$. ($T_c$ is the critical temperature.) Over the temperature and frequency ranges of the measurements [$10^{-3} < \tau < 10^{-1}$, 0.1 kHz$< f <$7.5 kHz], the attenuation was dominated by the thermal boundary layer. The model predicts that the attenuation at the boundary first increases as τ decreases and then saturates when the effusivity of the xenon exceeds that of the solid. [The effusivity is $\varepsilon \equiv \sqrt{\rho C_P \lambda_T}$, where C_P is the isobaric specific heat and λ_T is the thermal conductivity.] The model correctly predicts ($\pm 1.0\%$) the quality factors Q of resonances measured in a steel resonator ($\varepsilon_{ss} = 6400\ \mathrm{kg \cdot K^{-1} \cdot s^{-5/2}}$); it also predicts the observed increase of the Q, by up to a factor of 8, when the resonator is coated with a polymer ($\varepsilon_{pr} = 370\ \mathrm{kg \cdot K^{-1} \cdot s^{-5/2}}$). The thickness δ_T of the thermal boundary layer in the xenon decreases as τ decreases until $2\pi f \gamma \zeta/(\rho c^2) \approx 1$. ($\zeta$ is the bulk viscosity, γ is the heat capacity ratio, and c is the speed of sound.) For smaller τ, δ_T is predicted to become complex and increase. [Work supported by NASA.]

9:00

1aPAa5. The influence of gravity on Rayleigh streaming in thermoacoustic systems. Konstantin Matveev, Scott Backhaus, and Gregory W. Swift (Condensed Matter and Thermal Phys. Group, Los Alamos Natl. Lab., Los Alamos, NM 87545)

The thermal buffer tube is the component of thermoacoustic systems intended to transmit sound and to provide a thermal barrier between the hot and cold heat exchangers located at the opposite ends of the tube. In high-power operating regimes of thermoacoustic devices, the sound amplitude is high enough to cause significant mass (Rayleigh) streaming in the thermal buffer tube. The streaming flow convects heat, which can degrade the overall performance of a thermoacoustic system by thermally short-circuiting the hot and cold heat exchangers. This leads to increased heat consumption by prime movers and to reduced cooling capacity in refrigerators. The thermal buffer tube is characterized by a significant axial variation of temperature, so effects of gravity on the streaming should be important. A simple analytical model allows quantitative estimation of gravity effects, showing that gravity may significantly suppress the

streaming under certain conditions. The important parameters of the system and operating conditions that affect the streaming in the presence of gravity are identified. [Work supported by DOE's Office of Science.]

9:15

1aPAa6. An acoustic streaming instability within regenerator-based thermoacoustic devices. Jin H. So, Gregory W. Swift, and Scott Backhaus (Condensed Matter and Thermal Phys. Group, Los Alamos Natl. Lab., Los Alamos, NM 87545)

As traveling-wave thermoacoustic devices are scaled to higher power, the cross-sectional areas of their regenerators increase proportionally in order to maintain a thermodynamic optimum. Unfortunately, this allows the possibility of a spatially dependent acoustic streaming instability internal to their regenerators due to weak thermal and hydrodynamic communication transverse to the acoustic axis. Calculations show one possible result is two regions of opposed streaming forming a recirculating acoustic streaming cell similar to a convection cell. If the streaming is vigorous enough, it can transport enthalpy between heat exchangers at different temperatures, reducing the efficiency of the thermoacoustic device. An experimental apparatus is being used to search for this instability. A summary of the theory and results from the experimental search will be presented. [Work supported by DOE's Office of Science.]

9:30

1aPAa7. Maximum in the damping of shape oscillations of foam drops near the critical void fraction. Hai Wang, Li Liu, Gregory J. McDaniel, and Glynn R. Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, haiwang@bu.edu)

Small spheroidal samples of aqueous foam (''foam drops'') of varying gas volume fraction are acoustically levitated in an ultrasonic field. The normalized natural frequency and damping ratio are determined by treating a foam drop as a damped linear oscillator and measuring shape mode frequency response. The observed natural frequencies are compared to a wet foam model without any fitting parameters. Good agreement is found for gas volume fractions ranging from 0.01 to 0.87. The observed damping ratio of a foam drop exhibits a maximum for gas volume fractions 0.5 to 0.7. The data are inverted to infer the shear elastic modulus and its dependence on gas volume fraction for dry foams. [Work supported by NASA.]

9:45

1aPAa8. Magnetic resonance measurement of dynamics of cavitating fluid. Igor Mastikhin, Benedict Newling, Bruce J. Balcom, and Derrick Green (MRI Ctr., Phys. Dept., Univ. of New Brunswick, 8 Bailey Dr., Fredericton, NB E3B 5A3, Canada, mast@unb.ca)

The prevalent methods used in studies of cavitation are optical and acoustical. They are sensitive to changes in optical and/or acoustical transparency and are not applicable to studies of opaque media. Magnetic resonance (MR) methods, on the other hand, can be applied to arbitrarily opaque media, providing both dynamic and molecular information. A hindrance to implementation of MR methods in cavitation research is their relatively long measurement time: they cannot compete with optics or acoustics when a researcher needs a snapshot technique to study quickly changing processes on a microsecond scale. There are also some experimental problems with compatibility of MR scanners and acoustical equipment. However, if one is interested in longer time-scale processes, MR can provide data on velocity distribution, evaluate the turbulent processes in cavitating fluid, and measure correlations between the directions of flow on a microsecond scale. In this work, we show feasibility of an application of MR to studies of dynamics of cavitating fluid, with measurements of spatially resolved velocity spectra and other above-mentioned parameters.

10:00–10:15 Break

10:15

1aPAa9. BUSS: Final results from an experiment to determine the effect of gravity on sonoluminescence. Charles Thomas, Ronald Roy, and R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

The goal of the Boston University Sonoluminescence in Space (BUSS) project was to investigate the dependence of SBSL on ambient acceleration by performing experiments both on the KC-135 and in a land-based lab at Boston University. In this talk the final results of the KC-135 portion of the project will be discussed. Measurements of the maximum bubble Radius, bubble position, and peak light intensity emitted (R_{max}, z, and I_{peak}, respectively) were made during parabolic flight while keeping the acoustic pressure constant. Trends in R_{max}, z, and I_{peak} are studied by picking out their extrema from the data for each parabola. Additionally, the measured time-dependent R_{max}, z, and I_{peak} are compared to a quasi-static model which was developed to predict changes in these variables as a function of the changing acceleration and cabin pressure of the KC-135 during flight. Given the relatively good agreement of the model with measured data, it is concluded that the changes in R_{max}, z, and I_{peak} can be explained simply by considering the change in hydrostatic pressure caused by the variable acceleration. [Work supported by NASA.]

10:30

1aPAa10. Modulated single-bubble sonoluminescence: Dependence of phase of flashes, their intensity and rise/decay times on viscosity, the modulation strength, and frequency. Igor Mastikhin and Borko Djurkovic (MRI Ctr., Phys. Dept., Univ. of New Brunswick, 8 Bailey Dr., Fredericton, NB E3B 5A3, Canada, mast@unb.ca)

The single-bubble sonoluminescence (SBSL) signal was studied for the case of driving frequency modulated by lower frequency with an offset. In our work, the driving frequency of 28 kHz and the modulation frequencies of 25–1000 Hz were used. The modulation strength of 0.2, 0.5, and 0.8 was defined as the difference of highest and lowest pressures over modulation period. The measurements were performed for water–glycerol mixtures of various viscosities. The measured SBSL signal appeared as a train of flashes for modulation frequencies below 250 Hz, and as a continuous modulated signal for higher frequencies. At the same frequency, the flashes covered similar phase intervals for different modulation strengths and, accordingly, pressure ranges. At higher glycerol concentrations (up to 24%) both the intensity and the stability of flashes increased, due to damped shape instabilities and reduced dancing; however, the phase interval of flashes remained about the same. Such phase-locked behavior can be explained by translational movements of the bubble due to modulated Bjerknes force and changes in the symmetry of the bubble collapse. The changes in intensities and rise/decay times can serve as a measure of the gas exchange between the bubble and its surroundings during silent and luminescent intervals.

10:45

1aPAa11. Chemical oscillations in bubbles: Resolving the mystery of chaotic sonoluminescence. Charles Thomas, Ronald Roy, R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215), and Joachim Holzfuss (TU Darmstadt, 64289 Darmstadt, Germany)

One aspect of the Boston University Sonoluminescence in Space project included 1 g experiments to map the position of stable bubbles in (P_a,R_{max}) space, in water for a range of dissolved gas concentrations. While performing these experiments a regime was found in which R_{max} and the position of the bubble oscillate periodically on a slow timescale despite a *constant* acoustic pressure. Histograms of the phase of light emission from a bubble in such an oscillatory state reveal a broadband distribution. These measurements partially explain previous measurements

of quasiperiodic and chaotic sonoluminescence. We hypothesize that the oscillations are caused by the competing mass transfer mechanisms of rectified diffusion (growth) and chemical reaction (shrinkage). [Work supported by NASA.]

11:00

1aPAa12. Faraday film patterns. R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215) and John R. Saylor (Clemson Univ., Clemson, SC 29634-0921)

Locally flat liquid interfaces that are periodically forced to vibrate above a threshold amplitude exhibit an instability to waves of finite amplitude and half the forcing frequency. These waves (first observed and explored by Faraday) self-organize into a variety of patterns depending on experimental parameters. We report here the use of Faraday waves in liquid/particulate mixtures to cause deposition of the suspended particles in patterns that mimic the overlaying Faraday wave field. In this work the liquid/particulate mixture is deposited on a solid substrate in the form of a thin film. Hence the wave motion has a significant effect on the velocity field near the solid surface affecting the deposition locations of the particles. Subsequent evaporation of the liquid phase results in a stable, patterned, particulate film. We demonstrate this method experimentally at millimeter scales using relatively large particles, and discuss the possibility of using this method at higher frequencies to create particulate patterns at the micrometer and nanometer scales. Simulations of gold/silver particulate mixtures reveal variations in the particulate deposition pattern for these two particle types due solely to their density difference, suggesting a possible method for generating composite materials tailored on a very fine scale.

11:15

1aPAa13. Modeling of the bulk acoustic wave effects of electrode thin films in piezoelectric resonators. Sonal Srivastava and Yook-Kong Yong (Rutgers Univ., 623 Bowser Rd., Piscataway, NJ 08854)

The modeling of very high frequency piezoelectric resonators such as the thickness shear mode resonators is computationally and memorywise very intensive. The finite element mesh must be sufficiently fine so as to prevent the mesh impedance from interfering with the wave acoustics of the vibrating modes. The thin film of electrodes introduces additional complexity to the finite element models, hence an efficient method for treating the acoustic effects of such films are needed. Since the films on the resonator surfaces are stress-free, the film stiffness is small and negligible, and only the inertial effects need be modeled. Finite elements models are used to demonstrate the effects of electrode film stiffness and mass on the bulk acoustic wave characteristics of the piezoelectric plate. Results are presented to show the upper limits of the electrode film thickness upon which only the inertial effects are predominant. Results are also shown to demonstrate the efficiency of modeling only the inertial effects and neglecting the stiffness effects of thin electrode films.

MONDAY MORNING, 24 MAY 2004 NEW YORK BALLROOM B, 8:20 TO 11:30 A.M.

Session 1aPAb

Physical Acoustics: Recent Advances in Buried Landmine Detection I

James M. Sabatier, Cochair
University of Mississippi, National Center for Physical Acoustics, Coliseum Drive, University, Mississippi 38677

Bradley Libbey, Cochair
Night Vision and Electronic Sensors Directorate, 10221 Burbeck Road, Fort Belvoir, Virginia 22060

Invited Papers

8:20

1aPAb1. Nonlinear seismo-acoustic landmine detection. Dimitri Donskoy (Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030)

The seismo-acoustic methods are among the most promising emerging techniques for the detection of landmines. Numerous field tests have demonstrated that buried landmines manifest themselves at the surface through linear and nonlinear responses (or signatures) to acoustic/seismic excitation at the frequencies below 1000 Hz. The linear signatures are due to lower impedance of soil above softer mine, especially in the vicinity of mines resonance. The nonlinear signatures are explained by high contact nonlinearity at the mine–soil interface. These phenomena are utilized in high provability/low false alarm detection methods pioneered by University of Mississippi (linear detection) and Stevens Institute of Technology (nonlinear detection). A simple mass-spring model of the mine-soil system [Donskoy *et al.*, J. Acoust. Soc. Am. (2002)] explains and provides an analytical tool for analysis and prediction of both: linear and nonlinear signatures with respect to depths, soil, and mine types. This presentation provides an overview of theoretical and experimental investigations conducted at Stevens over the last 4 years with the emphasis on nonlinear detection techniques. Among major accomplishments are discovery and explanation of mines resonance behavior; soil/depth effect on buried mines resonances; discovery and analysis of nonlinear acoustic interactions at the soil–mine interface; and development of nonlinear quadratic and intermodulation detection algorithms based on dual-frequency excitation.

8:40

1aPAb2. Nonlinear acoustic experiments involving landmine detection: A connection between mesoscopic/nanoscale effects in geomaterials. Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

The vibration interaction between the top-plate interface of a buried plastic landmine and the soil above it appears to exhibit many characteristics of the mesoscopic/nanoscale nonlinear effects that are observed in geomaterials like rocks (sandstone) or granular materials. Experiments are performed with an inert VS 1.6 anti-tank mine that is buried 3.6 cm deep in dry sifted loess soil. Airborne sound at two primary frequencies $f_1 = 120$ Hz and $f_2 = 130$ Hz undergo acoustic-to-seismic coupling. Interactions with the compliant mine and soil generate combination frequencies that, through scattering, can affect the vibration velocity at the surface. Profiles of the soil surface particle velocity at f_1 and f_2 and the nonlinearly generated $f_1-(f_2-f_1)$ component are characterized by a single peak. Doubly peaked profiles at $2f_1+f_2$ and $2f_2+f_1$ are attributed to the familiar mode shape of a timpani drum. Near resonance, the bending (a softening) of a family of tuning curves for the soil surface vibration over a landmine exhibits a linear relationship between the peak frequency and the corresponding peak particle velocity, which also exhibit hysteresis effects. [Work supported by U.S. Army Communications-Electronics Command RDEC, NVESD, Fort Belvoir, VA.]

9:00

1aPAb3. Landmine detection using acoustic to seismic coupling: Theory and modeling. Roger Waxler (NCPA, Univ. of Mississippi, University, MS 38677, rwax@olemiss.edu) and Doru Velea (Planning Systems Inc., Reston, VA 20191)

It has been demonstrated that buried landmines can be found by insonifying the ground and then measuring the resulting vibration of the ground's surface. Over a buried landmine there is a large enhancement of the ground's vibration due to the excitation of mechanical modes of the mine. The resulting ground vibration can be understood as the near-field response to the resonant scattering of sound by the buried landmine. An effective fluid model has been developed. The results and limitations of this model will be discussed.

9:20

1aPAb4. Combining magnitude and phase features in acoustic landmine detection. Tsaipei Wang, James M. Keller (Dept. of Elec. and Computer Eng., Univ. of Missouri–Columbia, Columbia, MO 65211), and Paul D. Gader (Univ. of Florida, Gainesville, FL 32611)

The utility of acoustic-to-seismic coupling systems for landmine detection has been clearly established. They have been shown to be able to detect very low metal content landmines that are difficult to detect for ground-penetrating radars. This technique measures the difference in ground vibration velocity of regions with and without buried landmines when subject to acoustic excitation. For most applications, only the magnitude of the surface velocity is used to construct recognition algorithms. We recently introduced phase-based features in the classification scheme, significantly reducing false alarms at given detection probabilities. Here the focus is on the analysis of ground velocity data collected in the time domain with a moving array of laser Doppler vibrometers. The processing techniques used to extract the magnitude and phase information are described, as well as examples to demonstrate how combined magnitude and phase features help improve the detection of buried mines with weak signatures and reduce false alarms. Simple mass-and-spring models are shown to be useful in understanding the observed features. Also analyzed are the phase signatures of a number of ground regions with buried man-made clutter objects and how they can be useful in separating these clutters from actual landmine regions.

9:40

1aPAb5. Laser-induced acoustic landmine detection. Charles A. DiMarzio, Tianchen Shi, Florian J. Blonigen, and Stephen W. McKnight (Ctr. for Subsurface Sensing and Imaging Systems, Northeastern Univ., Boston, MA 02115)

Laser-induced acoustic (LIA) imaging is a new approach for underground object detection, especially for shallow buried landmine detection. This noncontact detection technique is based on a pulsed laser initiating photo-acoustic interaction in the ground. If sufficient photon energy couples into the ground, generating acoustic signals larger than seismic noise, a high enough mismatch of acoustic impedance between background media and target could provide a clear acoustic image revealing both the target's position and the 3-D shape information to reduce false alarms. In this paper, we present a model for laser-induced acoustic wave generation, including models of photo-acoustic sources and propagation of sound waves in media, to obtain a better understanding of basic physics behind the process. A photon-induced thermoelastic source and a photon-induced plasma source are investigated. Also, an investigation was performed in terms of attenuation, angular distribution, and phase velocities, and experimental results are reported and compared with theoretical analysis.

10:00–10:15 Break

10:15

1aPAb6. Seismic sonar for landmine detection and confirmation. Thomas G. Muir (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS, tmuir@olemiss.edu), Manell Zakharia, Aurore Grill, and Emanuel Gestat de Garambé (French Naval Acad., Brest Armees, France)

Impulsive vibration of the ground can generate seismic interface waves of the Rayleigh type. They decay exponentially with depth into the soil, and spread cylindrically with lateral range. At useful frequencies around 100 Hz, they typically travel at speeds around 100 m/s, with wavelengths around a meter. Rayleigh waves can be made to propagate in sonar-like pulses to buried targets, reflect, and return to the sonar for reception and signal processing, providing range, bearing, and information as to target type. We have conducted new experiments and analyses with seismic sonar in a clay soil. A focused array of ten sources and eight receivers (triaxial seismometers) were deployed at a range of 4.5 m to examine a 20-lb. landmine as well as a clump of rocks, and other false targets. After vector polarization processing, the amplitude of the mine target echo was 28 dB above the environmental backscatter. Mine-like target confirmation was provided by Wigner–Ville transformation, which allows separation of manmade from natural targets, by providing unique and identifiable time–frequency–amplitude signatures for each. The potential for target detection as well as a level of target type classification, at relatively long ranges, was demonstrated. [Work supported by the U.S. Army Night Vision Laboratory.]

10:30

1aPAb7. Demonstration of an end-fire array Rayleigh wave source for a seismo-acoustic sonar. Steven R. Baker and Steven E. Rumph (Phys. Dept., Naval Postgrad. School, Code PH/Ba, Monterey, CA 93943, srbaker@nps.navy.mil)

A linear array of four vertical-motion sources was deployed on the sand in the near-surf zone of Del Monte Beach, Monterey, CA. The sources were spaced 25 cm apart (approximately one-quarter wavelength at 100 Hz, the nominal operating frequency) and were driven with a transient signal in a sequential fashion so as to preferentially radiate Rayleigh waves in one end-fire direction. Beam patterns were measured at a radius of 3.5 m. Measurements were made with the array directing the radiation toward, away from, and parallel to the surf line. In general, results were in fair agreement with simple (nondispersive) theory, except for the depth of the nulls. A measured front-to-back radiation suppression of approximately 15 dB was routinely achievable.

10:45

1aPAb8. Land mine detection by time reversal acousto-seismic method. Alexander Sutin (Artann Labs., Inc. and Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030), Armen Sarvazyan (Artann Labs., Inc., Trenton, NJ), Paul Johnson, and James TenCate (Los Alamos Natl. Lab., Los Alamos, NM)

We present a concept and results of a pilot study on land mine detection based on the use of time reversal acoustics (TRA). TRA provides a possibility of highly effective concentrating of seismic wave energy in time and space in complex heterogeneous media. TRA focusing of seismic waves on a land mine increases the detection abilities of conventional linear and nonlinear acousto-seismic methods. Such factors as medium inhomogeneities, presence of reflecting boundaries, which could critically limit conventional acoustic approaches, do not affect TRA based method. The TRA mine detection system comprises several air borne or seismic sources and a noncontact (laser vibrometer) device for remote measurements of the surface vibration. The TRA system focuses a seismic wave at a surface point where the vibration is measured. The focusing point is scanned across the search area. The amplitude and frequency dependence of the signal from the seismic wave focusing point and nonlinear acoustic effects are analyzed to assess probability of the mine presence. Preliminary experiments confirmed high focusing ability of the TRA seismo-acoustic system in complex conditions (a laboratory tank with sand) and demonstrated a significant increase in the surface vibration in the presence of mine imitator. [Work supported by DoD grant.]

11:00

1aPAb9. Soil effect on landmine vibrations. Andrei Zagrai, Alexander Ekimov, and Dimitri Donskoy (Davidson Lab., Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030, azagrai@stevens.edu)

Field tests of the seismo-acoustic landmine detection techniques revealed strong dependence of mines' vibration signatures on the burial depth, soil physical properties, and environmental conditions. This study discusses the mines' linear and nonlinear signatures as a function of burial depth, soil type, and moisture content. Experimental results show that the resonance frequency of soil–mine system initially decreases with burial depth and then anomalously increases at greater depths. Soil moisture further amplifies this anomaly. Soil type and moisture content influence mines' nonlinear signatures as well. It was found that the quadratic nonlinear (Q) and cubic nonlinear (C) intermodulation mine vibrations respond differently to the presence of moisture in soil. While Q response diminished in wet soil (after rain), the C response was still highly evident and used for mine visualization. These observations revealed complex behavior of soil–mine system, which is not completely understood or explained. [Work supported by ONR.]

11:15

1aPAb10. Effects of elasticity and porosity in modeling of the acoustic-to-seismic transfer function. Margarita S. Fokina, Vladimir N. Fokin, and James M. Sabatier (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38655, mfok@olemiss.edu)

Modeling sound interaction with the ground is important both for remote sensing techniques and the elimination of false alarms in the landmine detection application. Elastic and porous-elastic models of the ground are most frequently used for these studies. Though both of these models are well known, few comparisons have been made between the acoustic-to-seismic transfer functions (TF) calculated from these two models for the frequency range 100–1000 Hz, which is typical for geoacoustic applications. In this work, the matrix technique was exploited to solve boundary equations for porous-elastic layers. This technique allows one to obtain the TF and refraction indexes for arbitrary porous-elastic stratifications of ground layers. The results of test computations of original codes are presented. Effects connected with the slow wave in the porous-elastic model are analyzed. Effective parameters of the visco-elastic model are analyzed to give the best fit to frequency dependence of the TF calculated in the frame of porous-elastic layered model. Comparison between real and effective parameters demonstrates the possible accuracy for obtaining properties of porous-elastic ground in the frame of elastic models. [Work supported by ONR Grant N00014-02-1-0878.]

Session 1aPPa

Psychological and Physiological Acoustics: Poster Session I (Poster Session)

Barbara Shinn-Cunningham, Chair
Cognitive and Neural Systems, Boston University, 677 Beacon Street, Boston, Massachusetts 02215

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

1aPPa1. Dissonance perception by listeners with sensorineural hearing loss. Jennifer B. Tufts, Marjorie R. Leek, and Michelle R. Molis (Army Audiol. and Speech Ctr., Walter Reed Army Medical Ctr., 6900 Georgia Ave. NW, Washington, DC 20307, jennifer.tufts@na.amedd.army.mil)

The perceived dissonance of two simultaneous tones (a dyad) depends upon their frequency separation (in the case of two pure tones) or their fundamental frequency ratio (in the case of two harmonic complex tones). The purpose of this study was to determine whether the perceived dissonance of puretone and harmonic complex dyads is altered in the presence of sensorineural hearing (SNHL), and, if so, whether this can be explained by the reduced frequency selectivity typically associated with SNHL. Four normal-hearing and four hearing-impaired listeners evaluated the dissonance of puretone and harmonic complex dyads centered at 500 and 2000 Hz. Frequency selectivity was estimated at 500 and 2000 Hz for each listener. People with SNHL rated the dissonance of puretone dyads similarly to the normal-hearing listeners, although auditory sensitivity and frequency selectivity differed considerably between the groups. However, their ratings of harmonic complex dyads did not show the pronounced differences in dissonance as a function of fundamental frequency ratio that were observed in the normal-hearing group. The poorer frequency selectivity of these listeners may have allowed more extensive interactions to occur among harmonic components, resulting in a less clear separation of dissonance and consonance. [Work supported by NIH-NIDCD.]

1aPPa2. Channel interaction in cochlear implants as a function of phase duration and pulse rate. Deniz Başkent and Robert V. Shannon (House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

Multichannel cochlear implants stimulate auditory nerves at different locations along the cochlea. It is widely assumed that good speech recognition requires independent activation of distinct tonotopic regions with no interactions between electrodes; interactions between electrodes can produce a result similar to spectral smearing. The electrical pulses produced by multiple electrodes can interact both spatially and temporally. However, little is known about the effect of basic electrical stimulation parameters on the degree of electrode interaction. It is possible that high stimulation pulse rates could improve the temporal representation of the signal but reduce the spatial (tonotopic) resolution due to increased electrode interaction. Differences in electrode stimulation configuration (monopolar or bipolar), pulse phase duration, pulse amplitude, and stimulation rate could change the level of electrode interaction. The present study used forward masking to measure electrode interaction patterns as a function of the stimulating pulse phase duration and the stimulation rate. A masker was placed on one electrode and the threshold elevation of a following signal was measured as a function of cochlear location. Forward-masked electrode interaction patterns will be presented for the different stimulation configurations and parameters for several masker levels. [Work supported by NIDCD Grant R01-DC-01526.]

1aPPa3. Place-pitch and excitation patterns in cochlear implant listeners. John J. Galvin III and Qian-Jie Fu (House Ear Inst., 2100 W. 3rd St., Los Angeles, CA 90057)

In cochlear implants (CIs), an electrode's pitch is largely determined by the excitation pattern produced by stimulation. For CI users, is pitch perception most strongly influenced by the peak, edge or some intermediate location within the excitation pattern? By varying the stimulation mode at four evenly spaced electrode locations, four spectral profiles were hypothesized: single-peaked, apically weighted, basally weighted or multi-peaked (with an intermediate pitch). Forward-masked excitation patterns and pitch judgments were obtained for the four electrode locations, for all experimental stimulation modes, relative to a standard set of BP1 electrodes. Results showed individual differences in CI users' excitation patterns and pitch judgments produced by the different stimulation modes. Most subjects' pitch judgments were sensitive to single, sharp peaks in the excitation pattern; however, subjects differed in their sensitivity to broader, multi-peaked patterns. For widely spaced electrode configurations, some subjects consistently judged pitch according to the apical edge of stimulation, some to the basal edge and others to an intermediate location. For some electrode locations, varying the stimulation mode produced significantly different pitches and excitation patterns. Future speech processing strategies may wish to combine stimulation modes to improve the spectral resolution available with a fixed number of implanted electrodes.

1aPPa4. Spectral profile discrimination ability using cochlear implants: Effects of number of active electrodes and pulse rate. Ward R. Drennan and Bryan E. Pfingst (Kresge Hearing Res. Inst., 1301 E. Ann St., Ann Arbor, MI 48109)

The ability of cochlear implant users to discriminate a change in spectral shape was investigated in listeners with Nucleus CI24M and CI24R(CS) implants. A primary independent variable was the number of active electrodes. Listeners were asked to detect a current increment to one of 3, 7, 11, and 21 active electrodes. Intensity discrimination on one electrode was also evaluated. Stimulation was achieved using a monopolar configuration with 75μs/phase biphasic pulses and a 24 μs interphase gap presented at 250 pps per channel. The pulses were swept in rapid succession from the base to the apex such that the overall rate increased with the number of active electrodes. Sensitivity to differences in spectral shape decreased with increasing number of electrodes. Next, the pulse rate was reduced to 159 pps per channel for the 11 and 21 active electrode conditions. Sensitivity to spectral shape at the slower pulse rate was better than

that at the higher rate given the same number of active electrodes. The improvement in sensitivity with the slower rate suggests that forward masking of one pulse over a successive pulse diminishes listener's ability to hear an intensity increment in the spectral profile.

1aPPa5. Psychophysical and speech results from the penetrating auditory brainstem implant (PABI). Robert V. Shannon, Mark E. Robert, and Steve Otto (House Ear Inst., 2100 W. Third St., Los Angeles, CA 90057)

The penetrating auditory brainstem implant (PABI) is a prosthetic device that delivers sound information to the brain via microstimulation of the cochlear nucleus. The design of the speech processor is similar to that in a cochlear implant. Surface electrodes are placed in the lateral recess of the IV ventricle, adjacent to the cochlear nucleus. Penetrating microelectrodes are inserted into the ventral cochlear nucleus. Threshold measures in two patients confirm intra-nuclear placement. Stimulation of the penetrating microelectrodes produced a full range of loudness sensations as well as a wide range of pitch sensations. Initial psychophysical measures show excellent temporal resolution (forward masking, gap detection, modulation detection) on the penetrating electrodes in one patient and poor temporal performance in another. Speech understanding was moderate, with one patient receiving primarily a supplement to lip-reading, while the other patient received modest open-set speech understanding on the third day following initial stimulation. Ongoing studies will characterize the interference between and within surface and penetrating electrode systems and the ability of the patient to integrate speech information across the two types of electrodes. [Work supported by NIDCD Contract N01-DC-00-11.]

1aPPa6. Amplitude modulation detection with cochlear implants: Effects of electrode separation and stimulus level. Anastasios Sarampalis and Monita Chatterjee (House Ear Inst., DAIP, 2100 W. Third St., Los Angeles, CA 90057)

Amplitude modulation (AM) detection performance has been studied in the past with normal-hearing and hearing-impaired populations. The temporal modulation transfer function (TMTF) is a plot of AM detection performance as a function of modulation rate and provides a way of characterizing temporal sensitivity. Typically the TMTF takes the form of a low-pass filter, with performance declining above 50–70-Hz modulation rate. TMTFs have also been measured with cochlear implant patients, showing a similar low-pass characteristic, with a cutoff around 140-Hz rate, while sensitivity to AM was found to increase with increasing current level. The present study investigated the effects of stimulation level and electrode separation on TMTFs with cochlear implant patients. TMTFs were measured for narrow through wide electrode separations and three different (loudness-balanced) percentages of the dynamic range. Preliminary results indicate that sensitivity increases (lower thresholds) with increasing stimulation level, for a given electrode separation. However, comparing TMTFs across different electrode separations, sensitivity is independent of current level, but increases as a function of percentage of dynamic range. In summary, it appears that AM detection performance with cochlear implants depends primarily on sensation level, rather than current level or electrode separation. [Work supported by NIDCD Grant No. R01DC04786.]

1aPPa7. Tactual temporal-onset order discrimination: Effects of frequency and site of stimulation. Hanfeng Yuan, Charlotte M. Reed, and Nathaniel I. Durlach (Res. Lab. of Electron., MIT, Cambridge, MA 02139)

This research is concerned with measurements of temporal onset-order discrimination through the tactual sensory system. Sinusoidal signals were delivered through a multi-finger stimulating device that operates over a frequency range of roughly 0 to 300 Hz and an amplitude range of roughly 0 to 50 dB SL. Measurements were obtained using a one-interval two-alternative forced-choice procedure in which each interval consisted of the random-order presentation of two different sinusoidal signals whose amplitude and duration were roved independently from trial to trial. Thresholds were estimated from psychometric functions of d' as a function of stimulus-onset asynchrony. Performance was studied as a function of the frequency separation between the two sinusoids presented in a given run for presentation across two different fingers and for presentation to the same finger. Conditions were also included to examine the effects of redundancy of frequency and site of stimulation on performance. Results will be compared across conditions and will be discussed with regard to their implications for the design of tactual displays of speech for persons with profound hearing impairment. [Research supported by Grant No. 5 R01-DC00126 from NIDCD, NIH.]

1aPPa8. Monaural informational masking release in children and adults. Emily Buss, Joseph W. Hall III, and John H. Grose (Univ. of North Carolina at Chapel Hill, 130 Mason Farm Rd., CB7070, 1115 Bioinformatics Bldg., Chapel Hill, NC 27599, ebuss@med.unc.edu)

Informational masking refers to an elevation in signal threshold due to stimulus uncertainty, rather than to energetic masking. This study assessed informational masking and utilization of cues to reduce that masking in children aged 5–9 and adults. We used a manipulation introduced by Kidd *et al.* [J. Acoust. Soc. Am. **95**, 3475–3480 (1994)] in which the signal was a train of eight consecutive tone bursts, each at 1 kHz and 60 ms in duration. Maskers were comprised of a pair of synchronous tone-burst trains whose frequencies were selected from the range spanning 0.2–5 kHz, with a protected region 851–1175 Hz. In the reference condition, where informational masking is pronounced, these maskers were eight bursts and had a fixed frequency within each interval, with new frequencies chosen randomly prior to each interval. Two conditions of masking release were tested: random frequency selection for each masker burst and a masker leading fringe of two additional 60-ms bursts. Both children and adults showed a significant informational masking effect, with children showing a larger effect. Both groups also showed significant release from masking, though initial results suggest that this may have been reduced in the youngest children. [Work supported by NIH, RO1 DC00397.]

1aPPa9. Differential effectiveness of cues in informational masking studies. Virginia M. Richards, Rong Huang (Dept. of Psych., Univ. of Pennsylvania, Ste. 302C, 3401 Walnut St., Philadelphia, PA 19104, richards@psych.upenn.edu), and Gerald Kidd, Jr. (Boston Univ., Boston, MA 02215)

For the detection of a tone added to random multitone maskers, playing a preview of the masker before detection trials can reduce thresholds compared to when there is no preview. In contrast, playing a preview of the signal-plus-masker does not provide a release from masking. This differential effectiveness of cues was examined in several conditions. Using the method of constant stimuli and a yes/no task, observers detected a 1000-Hz tone added to six-tone maskers. Prior to each trial, the frequencies of the masker components were randomly drawn. Two types of cues were tested, either a copy of the masker or a copy of the signal-plus-masker. The cues were presented either before or after the yes/no presentation interval. Finally, data were collected either blocked for each condition or the trials from the four conditions were interleaved. *D*-prime values were higher when the conditions were blocked than when they were interleaved. The pattern of results was the same in both situations; *d* was highest for pretrial masker cues, lowest for pretrial signal-plus-masker cues, and intermediate when either cue followed the trial interval. [Work supported by NIH/NIDCD.]

1aPPa10. Different learning patterns for tone detection in three simultaneous-masking conditions with different temporal characteristics. Julia L. Huyck and Beverly A. Wright (Dept. of Commun. Sci. and Disord. and Northwestern Univ. Inst. for Neurosci., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60201-3550)

Can the processes underlying performance on basic masking tasks be modified with practice? To address this question, 17 listeners completed pre- and posttests in which detection thresholds for a 20-ms 1-kHz tonal signal were measured in three notched-noise simultaneous-masking conditions that differed in the masker duration and relative onset times of the signal and masker, and three forward-masking conditions that differed in the tonal frequencies of the maskers. For the ~8 days between tests, eight of these listeners practiced 240 trials/day on each simultaneous-masking condition. Learning differed across conditions. When the signal and a 20-ms masker began and ended together, both trained and untrained listeners improved, but trained listeners improved more, learning gradually. When the signal began 400 ms after a 500-ms masker, only the trained group improved, learning quickly. When the signal and the 500-ms masker began together, both groups improved equally. Both groups also improved equally on each of the three forward-masking conditions. These results suggest that processes underlying performance on simultaneous- and forward-masking tasks are modifiable, training affects different mechanisms in simultaneous-masking conditions with different temporal characteristics, and the simultaneous-masking mechanisms affected by multihour training are not involved in forward masking. [Work supported by NIH.]

1aPPa11. Training experienced hearing-aid users to identify syllable-initial consonants in quiet and noise. James D. Miller, Jonathan M. Dalby, Charles S. Watson, and Deborah F. Burleson (Commun. Disord. Technol., Inc., Indiana Univ. Res. Park, 501 N. Morton St., Bloomington, IN 47404, jdmiller@artsci.wustl.edu)

Five experienced hearing-aid users with sensorineural hearing loss were given 14 h of intensive training identifying consonants in quiet and noise. Their performance was compared to that of five similar hearing-aid users with no special training. All listeners had moderate to severe hearing losses and had worn hearing aids for at least 1 year. All were pretested with a set of 20 consonants combined with three vowels /I,a,u/ as spoken by six different talkers. Pretests were conducted in quiet and in noise (multitalker babble) at moderate signal-to-noise ratios (SNRs). Training was conducted with eight target consonants (TCs). The TCs were in each listener's middle range of difficulty, and the three most common confusors for each target were individually selected, forming target sets of four consonants. Training was conducted in quiet and noise. During training, trial-by-trial feedback was given and, following an error, the listener could rapidly compare the intended syllable with its confusor. In noise, the SNR adapted to a criterion of 80 correct. There were no differences between training and control listeners on the pretests. After training, there was a significant 5% advantage for the trained listeners. Training generalized to talkers never heard during training. [Work supported by NIDCD.]

1aPPa12. Evidence for auditory signal-detection templates tuned to both the frequency and duration of an expected signal. Beverly A. Wright (Dept. of Commun Sci. and Disord. and Inst. for Neurosci., 2240 Campus Dr., Northwestern Univ., Evanston, IL 60208-3550)

When trying to detect a tonal signal in a continuous broadband noise, listeners attend selectively to both the frequency and the duration of the expected signal. However, it is not known whether they monitor separate, or combined, representations of these two attributes. To investigate this question, a probe-signal method was used to measure the detectability of signals of expected and unexpected durations at two expected frequencies. Four listeners were led to expect one of two signals to be presented at random: a brief tone at one frequency or a long tone at another frequency. For each signal frequency, the detectability of the signals of unexpected duration decreased to near chance as the difference between the expected and unexpected duration, at that frequency, increased. Thus, signals of each expected duration were rarely detected when they were presented at a frequency not associated with that duration. The frequency specificity of this duration tuning indicates that both the frequency and the duration of an expected stimulus are represented in a single signal-detection template. [Work supported by NIH.]

1aPPa13. ZEST as a tool for rapid assessment of frequency discrimination. Michael Stahl, Jr., Søren Buus (Inst. for Hearing, Speech, and Lang. and Commun. and DSP Ctr., 440 DA, Northeastern Univ., 360 Huntington Ave., Boston, MA 02215), and Mary Florentine (Northeastern Univ., Boston, MA 02215)

The purpose of this study is to develop a rapid and reliable procedure for obtaining discrimination thresholds. ZEST [King-Smith *et al.*, Vision Res. **34**, 885–912 (1994)] has been found to be a promising candidate for this purpose. The present study used ZEST to obtain frequency-discrimination thresholds in 30 trials with a 2AFC paradigm. Subsequent analysis allowed calculation of thresholds for any number of trials up to 30. The stimuli were 600-ms tones at seven frequencies, ranging from 250 to 7000 Hz. Data for six normal listeners obtained with three different slopes (beta) of the assumed psychometric function indicate that reliable thresholds can be obtained in 10 to 15 trials. Simulations indicate that threshold estimates have only slight bias ranging from 15% with 9 trials to under 5% with 30 trials. Interestingly, these simulations also show that this bias can be reduced with only a small increase in the variability if the assumed psychometric function used by ZEST is made steeper than the listener's psychometric function (i.e., if beta is increased from 1.23 to 6). Altogether these results suggest ZEST combined with a 2AFC paradigm is a promising candidate for rapid and reliable assessment of listeners' discrimination thresholds.

1aPPa14. The effect of compression and attention allocation on speech intelligibility. II. Sangsook Choi and Thomas Carrell (Dept. of Special Ed. and Commun. Disord., Univ. of Nebraska—Lincoln, 253 Barkley Memorial Ctr., Lincoln, NE 68583-0731, schoi6@bigred.unl.edu)

Previous investigations of the effects of amplitude compression on measures of speech intelligibility have shown inconsistent results. Recently, a novel paradigm was used to investigate the possibility of more consistent findings with a measure of speech perception that is not based entirely on intelligibility (Choi and Carrell, 2003). That study exploited a dual-task paradigm using a pursuit rotor online visual-motor tracking task (Dlhopolsky, 2000) along with a word repetition task. Intensity-compressed words caused reduced performance on the tracking task as compared to uncompressed words when subjects engaged in a simultaneous word repetition task. This suggested an increased cognitive load when listeners processed compressed words. A stronger result might be obtained if a single resource (linguistic) is required rather than two (linguistic and visual-motor) resources. In the present experiment a visual lexical decision task and an auditory word repetition task were used. The visual stimuli for the lexical decision task were blurred and presented in a noise background. The compressed and uncompressed words for repetition were placed in speech-shaped noise. Participants with normal hearing and vision conducted word repetition and lexical decision tasks both independently and simultaneously. The pattern of results is discussed and compared to the previous study.

1aPPa15. Thresholds for inattentive listeners obtained with adaptive forced-choice procedures. Robert Schlauch and Edward Carney (Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Forced-choice adaptive procedures enjoy widespread use for measurement of detection and discrimination thresholds. Using computer simulations and behavioral data, this paper examines the effect of step size (2 and

4 dB), decision rule (targets =70.7%, 75%, and 79.4% correct) and initial starting level on thresholds obtained under conditions of inattention and complete attention. Simulations of two-alternative forced-choice (AFC) and three-AFC procedures demonstrate that the best general strategy for measuring thresholds under conditions of inattention is to employ a low percentage (e.g., 70.7%) as a target on the psychometric function combined with a large step size. When a track for an inattentive listener is begun far above threshold, we found that using a large step size for the first two turnarounds (reversals in stimulus level direction) and then switching to a step size that is $\frac{1}{2}$ the initial step size was less effective than using the large step size for the entire threshold estimation run. Behavioral data obtained in a tonal detection-in-noise task with a 2-AFC procedure support the validity of the simulations.

1aPPa16. Effects of N-acetylcysteine on noise-induced temporary threshold shift and temporary emission shift. Martin Robinette (U.S. Army Aeromedical Res. Lab., 6901 Andrews Ave., Fort Rucker, AL 36362-0577, martin.robinette@us.army.mil)

Animal research has shown that antioxidants can provide significant protection to the cochlea from traumatic noise exposure with some benefit when given after the exposure. Similar results in humans would have a significant impact on both prevention and treatment of noise-induced hearing loss. The current study evaluates the effectiveness of N-acetylcysteine (NAC) on temporary threshold shift (TTS) by using both behavioral and physiological measures. Sixteen healthy, normal-hearing subjects were given NAC or a placebo prior to exposure to a 10-min, 102-dB narrowband noise, centered at 2 kHz. This exposure was designed to induce a 10–15-dB TTS. Following the noise exposure, pure-tone thresholds (Bekesy) and transient-evoked otoacoustic emissions (TEOAE) were measured for 60 min to monitor the effects of NAC on TTS recovery. Postexposure measures were compared to baseline data. [Work supported by American BioHealth Group.]

1aPPa17. Acute effects of nicotine on processing of complex stimuli in smokers and nonsmokers. Ashley Harkrider and Mark Hedrick (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 South Stadium Hall, Knoxville, TN 37996, aharkrid@utk.edu)

Effects of nicotine in the auditory system of normal-hearing smokers and nonsmokers were investigated both behaviorally and physiologically. Discrimination of consonant–vowel speech in quiet and noise was assessed in the presence and absence of a transdermal nicotine patch by measuring categorical boundaries and mismatch negativity (MMN). Data indicate that the effects of nicotine on both behavioral and physiological measures increased with an increase in severity of nicotine-induced symptoms. Smokers showed improved CV discrimination in quiet and noise with nicotine. Additionally, smokers exhibited more measurable and significantly sharper boundaries as well as larger MMN areas than nonsmokers in quiet and noise for both placebo and nicotine sessions. MMN data acquired for both quiet and noise, and behavioral data acquired in quiet, indicate that smokers show the greatest improvements in discrimination during nicotine exposure, followed by symptomatic nonsmokers. Asymptomatic nonsmokers show little improvement with nicotine and, on occasion, show decrements in performance. These data may contribute to our understanding of the role of nAChRs in the auditory system, the neural mechanisms that underlie the recognition of sound in quiet and noise, and mechanisms mediating improved information processing and enhanced cognitive performance that serve as reinforcement for continued tobacco use by smokers.

1aPPa18. Fractal EEG analysis with Higuchi's algorithm of low-frequency noise exposition on humans. Ryszard Panuszka, Zbigniew Damijan, and Cezary Kasprzak (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Eng. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland, damijan@imir.agh.edu.pl)

Authors used methods based on fractal analysis of EEG signal to assess the influence of low-frequency sound field on the human brain electro-potentials. The relations between LFN (low-frequency noise) and change in fractal dimension EEG signal were measured with stimulations tones. Three types of LFN stimuli were presented; each specified dominant frequency and sound-pressure levels (7 Hz at 120 dB, 18 Hz at 120 dB, and 40 Hz at 110 dB). Standard EEG signal was recorded before, during, and after subject's exposure for 35 min. LFN. Applied to the analysis fractal dimension of EEG-signal Higuchis algorithm. Experiments show LFN influence on complexity of EEG-signal with calculated Higuchi's algorithm. Observed increase of mean value of Higuchi's fractal dimension during exposition to LFN.

1aPPa19. Low-frequency sounds and psychological tests at 7, 18, and 40 Hz. Zbigniew Damijan, Cezary Kasprzak, and Ryszard Panuszka (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Eng. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland, damijan@imir.agh.edu.pl)

Research included the results of tests aimed to determine how LFN (low-frequency noise) with the dominating frequency (7 Hz at 120 dB, 18 Hz at 120 dB, and 40 Hz at 110 dB) influences human brain potentials and understanding the dependency of results achieved in psychological questionnaires. The psychological questionnaires (EPQ-R Eysencks and SSS-5 Zuckermans) were analyzed. Presented issue existences difference, relative influence LFN on human biopotentials, dependence from acquired results in ranges EPQ-R Eysencks and SSS-5 Zuckermans. The test included 96 experiments. Standard EEG potentials, ECG potentials, and EDP (dermal) were recorded before, during and after subject's 35-min exposures to LFN. Evident differences in changed bio-signals especially in EEG subject's dispersion, were easily determined and correlated to questionnaire reports.

1aPPa20. Sound fields in biosphere of the mountain streams and their influence on the human EEG. Cezary Kasprzak, Zbigniew Damijan, and Ryszard Panuszka (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland, cekasp@agh.edu.pl)

Low-frequency acoustics fields in the biosphere of the headwater regions in Poland (Bieszczady Mountains and Tatry Mountains) were researched and the the influence of observed low-frequency acoustics field on changes of human EEG signals was analyzed. Several places were found with specific distribution of sound fields where low frequencies dominated. Interesting parameters of distribution of the riverbeds and physical characterization of flows were determined. Standard EEG signal was recorded from subjects until exposure for 20 min. It was ascertained that spectra of sound-pressure levels were with specific shape, very similar for a group of rivers. The influences of SPL on level of power spectrum of EEG signal were analyzed. Changes in the brain waves were observed with increases in delta and decreases in alpha. Tested subjects reported behavior specific for relaxation.

Session 1aPPb

Psychological and Physiological Acoustics: Pitch

Andrew J. Oxenham, Chair

Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Contributed Papers

8:45

1aPPb1. Effects of stimulus level and harmonic resolvability on pitch discrimination. Joshua G. W. Bernstein and Andrew J. Oxenham (MIT Res. Lab. of Electron. and Harvard-MIT Speech & Hearing Bioscience & Technol. Prog., 77 Massachusetts Ave., Cambridge, MA 02139, jgbern@mit.edu

Fundamental frequency ($F0$) discrimination performance depends largely on harmonic number: only complexes containing harmonics below the 10th yield small $F0$ difference limens (DLs). This may be because only the low harmonics are resolved by the auditory periphery. Auditory filter bandwidths are known to increase with increasing stimulus level, providing a tool for testing the effect of reduced harmonic resolvability on $F0$ discrimination. $F0$DLs were measured for equal-amplitude, random-phase harmonic complexes, with eight different $F0$s between 75 and 400 Hz, bandpass filtered between 1.5 and 3.5 kHz. Complexes were presented at a constant average sensation level (12.5 dB/component before filtering) in threshold equalizing noise (TEN) with levels of 10, 40, and 65 dB SPL per equivalent rectangular bandwidth (ERBn). If good $F0$ discrimination requires the presence of resolved harmonics, the transition from poor to good $F0$ discrimination with increasing $F0$ should shift to higher $F0$s at higher levels, because fewer harmonics will be resolved. Preliminary results indicate that three out of four subjects demonstrated this shift at the highest level tested, providing some support for the idea that peripherally resolved harmonics are required for good pitch discrimination performance. [Work supported by NIH Grants R01DC05216 and 5T32DC00038.]

9:00

1aPPb2. Detection and $F0$ discrimination of harmonic complex tones in the presence of concurrent complexes. Christophe Micheyl and Andrew Oxenham (Res. Lab. of Electron., MIT, Bldg. 36-758, Cambridge, MA 02139-4307, cmicheyl@MIT.EDU)

This study measured the level of target harmonic complex tones, relative to simultaneous concurrent harmonic or inharmonic maskers, necessary for (a) target detection and (b) discrimination of the target $F0$. Depending on the condition, the $F0$ of the target was around 100, 200, or 400 Hz and the average $F0$ of the masker was either equal to this or 7 semitones lower or higher. Target and maskers were bandpass filtered between 1200 and 3600 Hz (48-dB/octave slopes). Low-pass noise masked distortion products. Masked thresholds in the target-detection task were generally higher when the average $F0$ of the masker was lower than that of the target, and were also higher for inharmonic than for harmonic maskers. In the $F0$ discrimination task, average $F0$ differences between the target and masker tones consistently improved masked thresholds at the highest $F0$ used (400 Hz) but not the lowest (100 Hz). Temporal asynchronies between the target and masker and reduced masker–$F0$ uncertainty had similar effects. These results suggest that peripheral frequency resolution may limit the ability to benefit from auditory segregation cues for both the detection and discrimination of harmonic complex tones embedded in other tones. [Work supported by NIDCD R01DC05216.]

9:15

1aPPb3. Dominance region for pitch: Effect of duration. Hedwig Gockel, Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, 15 Chaucer Rd., Cambridge CB2 2EF, UK, hedwig.gockel@mrc-cbu.cam.ac.uk), and Christopher J. Plack (Univ. of Essex, Colchester CO4 3SQ, UK)

Frequency discrimination thresholds for pure tones increase with decreasing duration, more so at low than at high frequencies. The present study investigated the possible influence of this frequency-dependent effect on the dominance region for pitch (DRP) of complex tones (CTs). The DRP was determined for CTs containing the first seven harmonics with a fundamental frequency ($F0$) of 250 Hz, and a level of 65 dB SPL per component. The tones had a duration of 16 or 200 ms, including 8-ms ramps. Listeners were presented with a CT with one of the harmonics mistuned upwards or downwards by 3%, followed 500 ms later by a perfectly harmonic CT of the same duration. Listeners had to adjust the $F0$ of the harmonic tone so that its pitch matched that of the mistuned complex. For all five listeners, the DRP was higher in harmonic number for the short than for the long duration tones. The results are consistent with the idea that pitch is derived from a weighted sum of information from individual harmonics, the individual weights being dependent on duration and related to the discriminability of the harmonics. [Work supported by EPSRC Grant GR/R65794/01.]

9:30

1aPPb4. Perceiving a change in a nonperceived pitch. Laurent Demany and Christophe Ramos (CNRS and Univ. Victor Segalen, 146 rue Leo Saignat, F-33076 Bordeaux, France)

Listeners were presented with 300-ms "chords" of five synchronous pure tones, followed after a 0.5–8-s silence by a single pure tone. The frequencies of each chord's components were selected randomly, but spaced by intervals of between 6 and 10 semitones. In one condition ("up/down"), the single tone following a chord was 1 semitone higher or lower than one of the chord's three intermediate components; on each trial, the corresponding component was selected randomly and the task was to indicate the direction in which its pitch changed. In another condition ("present/absent"), the single tone following a chord was either identical to one of the three intermediate components or halfway in frequency between two components; the task was to indicate if the single tone was present in the chord or not. Performance was much better in the up/down condition than in the present/absent condition, even though the opposite trend was predictable for an ideal "analytic" listener. Ten listeners reported that, in the up/down condition, they could often perceive the appropriate pitch change without having heard out the relevant component of the chord. These results provide strong evidence for the existence of pitch change detectors in the auditory system.

9:45

1aPPb5. Fundamental frequency discrimination for resolved and unresolved harmonics with the same pitch in the same spectral region. Rebecca K. Watkinson, Christopher J. Plack (Dept. of Psych., Univ. of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK, rwatki@essex.ac.uk), and Robert P. Carlyon (MRC Cognition and Brain Sci. Unit, Cambridge, U.K.)

To investigate the hypothesis that fundamental frequency ($F0$) is encoded via different mechanisms for resolved and unresolved harmonics, $F0$ discrimination was measured for harmonics that had the same pitch and were in the same spectral region, but which differed in resolvability. Summing unresolved harmonics in alternating phase increases their pitch by an octave relative to sine-phase summation; this is not found for resolved harmonics. An alternating-phase, 100-Hz $F0$ complex tone and a sine-phase, 200-Hz $F0$ complex tone, both filtered between 1000 and 2000 Hz, have the same pitch. However, the first complex tone contains unresolved harmonics and the second contains resolved harmonics. The experiment measured sequential $F0$DLs and d's for two such groups of harmonics that were both resolved, both unresolved, or for which one group was resolved and the other unresolved. Performance was worse for unresolved versus unresolved comparisons than for resolved versus resolved comparisons. More importantly, performance was worse still when the harmonic groups differed in resolvability (resolved versus unresolved). This provides some evidence for the hypothesis that resolved versus unresolved comparisons are impaired by the need to compare the outputs of separate pitch mechanisms. [Work supported by an EPSRC doctoral training grant.]

10:00–10:15 Break

10:15

1aPPb6. The role of timbre in pitch matching abilities and pitch discrimination abilities with complex tones. Robert E. Moore (Dept. of Speech Pathol. and Audiol., Univ. of South Alabama, 2000 UCOM, Mobile, AL 36688-0002, rmoore@usouthal.edu), Christopher R. Watts, and Fawen Zhang (Univ. of South Alabama, Mobile, AL 36688-0002)

Control of fundamental frequency (F_0) is important for singing in-tune and is an important factor related to the perception of a talented singing voice. One purpose of the present study was to investigate the relationship between pitch-matching skills, which is one method of testing F_0 control, and pitch discrimination skills. It was observed that there was a relationship between pitch matching abilities and pitch discrimination abilities. Those subjects that were accurate pitch matchers were also accurate pitch discriminators (and vice versa). Further, timbre differences appeared to play a role in pitch discrimination accuracy. A second part of the study investigated the effect of timbre on speech discrimination. To study this, all but the first five harmonics of complex tones with different timbre were removed for the pitch discrimination task, thus making the tones more similar in timbre. Under this condition no difference was found between the pitch discrimination abilities of those who were accurate pitch matchers and those who were inaccurate pitch matchers. The results suggest that accurate F_0 control is at least partially dependent on pitch discrimination abilities, and timbre appears to play an important role in differences in pitch discrimination ability.

10:30

1aPPb7. Establishing a hierarchy among octave equivalent forms of a 12-tone row. Hubert C. Ho (Ctr. for New Music and Audio Technol., Univ. of California, Berkeley, 1750 Arch St., Berkeley, CA 94720, hubertho@uclink.berkeley.edu)

The perceptability of tone hierachies in dodecaphonic rows has been well documented in the literature. This study investigates similarity among octave-equivalent variants of a given dodecaphonic row. The aggregate pitch proximity of a given row is represented by a 12-dimensional "contour vector," the elements of which consist of the intervals (in semitones) of successive members of the row. Participants were asked to perform a series of probe-tone rating tasks in a variety of given contexts according to techniques used by Krumhansl *et al.* (1987). Test rows varied with respect to the contour vector. Control rows consisted of randomly generated tone rows. It was found that the similarity between contour vectors of a given row is positively correlated with their respective probe tone ratings.

10:45

1aPPb8. Low pitches of frequency-transposed stimuli: Salience estimates from a full population-interval model. Peter Cariani (Eaton Peabody Lab., Mass. Eye and Ear Infirmary, 243 Charles St., Boston, MA 02114 cariani@epl.meei.harvard.edu)

Transposed stimuli consisting of half-wave rectified modulators of high frequency carriers allow comparisons of pitch perception across frequency regions. While listeners can discriminate low $F0$-pitches of simple transposed stimuli (1 Fm-Fc), most cannot detect low pitches of complex transposed stimuli [Fm=300, 400, 500 Hz, Fc=4, 6.35, 10.08 kHz; Oxenham *et al.*, PNAS (2004)]. Auditory nerve (AN) responses to a low-frequency harmonic complex ($F0=100$ Hz;$n=3–5$) and its transposition were simulated using filtering (middle ear, 48 gammatone filters, 144 fibers, human CF distribution, 100 Hz to 18 kHz), rectification, adaptive gain control, spontaneous activity (three classes), 65 dB SPL simulated level. Pitch salience was estimated by analyzing population-interval histograms using subharmonic interval sieves. The harmonic complex produced high salience (2.05), well above assumed pitch detection threshold ($\sim$1.3). Simple transposed stimuli produced lower saliences (1–1.37). Some were below threshold, inconsistent with psychophysical pitch discrimination results. The complex transposed stimulus produced an $F0$-salience (1.31) near threshold, qualitatively consistent with their mixed detection results. Dramatic differences are thus seen between low-frequency and transposed stimuli. Clearly, for these and other high-frequency stimuli, full AN population-interval models can produce salience estimates that differ substantially from simple intuitions and/or partial implementations. Interpretational caution is therefore counseled.

11:00

1aPPb9. Fundamental frequency discriminability and utility in normal-hearing listeners using noise-excited vocoders. Michael Qin and Andrew Oxenham (MIT Res. Lab. of Electron. and Harvard-MIT Div. of Health Sci. and Technol., SHBT Prog., 77 Massachusetts Ave., Cambridge, MA 02139-4307, qin@mit.edu)

Although enormous progress has been made in the development of cochlear implants, the speech reception of implant users is still not comparable to that of normal-hearing listeners. The difference in performance is especially pronounced in complex auditory situations, such as in the presence of competing talkers. The poorer speech reception performance of implant users in complex environments may, at least in part, be due to the poorer fundamental frequency ($F0$) representation. The present study examined the effects of implant-like processing (i.e., noise-excited vocoding) on $F0$ discriminability and utility in normal-hearing listeners. $F0$ difference limens ($F0$DLs) were measured as a function of the number of vocoder channels. In addition, vowel identification was measured as a function of $F0$ difference between competing vowels in a double-vowel paradigm, with number of vocoder channels as a parameter. Our findings show that despite the reasonable $F0$DLs (less than one semitone) with 24- and 8-channel vocoder processing, listeners were unable to benefit from $F0$ differences between the competing vowels in a double-vowel paradigm. The implications of the findings for pitch theories and cochlear implant design will be discussed. [Work supported by NIDCD Grant R01 DC05216.]

Session 1aSC

Speech Communication: Poster Session I

Susan Shaiman, Chair
Department of Communication Sciences and Disorders, University of Pittsburgh, 4033 Forbes Tower, Pittsburgh, Pennsylvania 15260

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

1aSC1. Evaluating the effects of bilingual dominance and language mode on overall degree of foreign accent. Satomi Imai, James E. Flege (Div. of Speech and Hearing Sci., Univ. of Alabama at Birmingham, Birmingham, AL 35294), and Ian R. A. MacKay (Univ. of Ottawa, Ottawa, ON, Canada)

This study evaluated the influence of bilingual dominance and language mode on overall degree of perceived foreign accent. Three groups of Italian–English bilinguals ($n=12$ each) were selected according to their ratio of self-rated English/Italian proficiency: English-dominant, balanced, or Italian-dominant. Language mode was manipulated by having participants repeat English phrases before and after similar Italian phrases (E1, E2) and then intermixed with Italian phrases (E3). Native English (NE) listeners rated four English phrases spoken by the bilinguals and 12 age-matched NE controls using a scale that ranged from 1 (strong foreign accent) to 9 (no foreign accent). We hypothesized that if switching into the native language (here, Italian) adversely affects pronunciation of the second language (English), the third repetitions of the English phrases (E3) should be more strongly foreign-accented than the first repetitions (E1). The foreign accent ratings decreased significantly in the following order: NE > English-dominant > balanced > Italian-dominant. That is, all three bilingual groups had detectable foreign accents, and strength of accent depended on bilingual dominance. The language mode effect was significant only for one of the four phrases examined, perhaps because it (*mozzarella cheese*) has distinctly different phonetic renditions in English and Italian. [Work supported by NIH.]

1aSC2. A cross-language study on perception of Taiwanese stops by non-native listeners. Yueh-chin Chang (Natl. Tsing Hua Univ., 101, Sec. II, Kuang Fu Rd., Hsinchu, Taiwan), Catherine T. Best (Wesleyan Univ., Middletown, CT 06459), and Pierre A. Halle (CNRS-Paris V, 92774 Boulogne-Billancourt, France)

Few reports exist on perception of three-way stop voicing distinctions by non-native listeners whose languages have two-way distinctions that vary phonologically and phonetically. We examined perception of Taiwanese stops by French, Mandarin, and American English listeners. Taiwanese has three voicing categories: unaspirated voiceless (U), aspirated voiceless (A) and voiced (V). Phonologically, English and French have voiced-voiceless contrasts, so these listeners should have similar difficulties with Taiwanese /U-A/. Phonetically, however, English and Mandarin distinguish voiceless unaspirated versus aspirated stops, in which case these listeners should have equivalent difficulties with Taiwanese /U-V/. American and Mandarin listeners' discrimination supported the second prediction. French listeners discriminated /U-A/ better than /U-V/ for velar stops, but the opposite for labials, possibly because French velar stop VOTs preceding high vowels are longer than Mandarin and American ones. Mandarin listeners discriminated better than French listeners overall. American listeners' discrimination was poorest. The discrimination results are consistent with identification patterns for the three groups. The findings suggest that speakers of languages with phonologically unaspirated versus aspirated contrasts (Mandarin) can distinguish three voicing types more easily than speakers of languages with a voiced versus voiceless contrast (French, English), especially when the phonetic realizations differ from the phonological distinction (English).

1aSC3. Production and perception of a temporal contrast by native and non-native speakers. Tessa Bent, Ann R. Bradlow (Dept. of Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208, t-bent@northwestern.edu), and Bruce Smith (Univ. of Utah, Salt Lake City, UT 84112)

In previous work, we found that both native and Chinese-accented talkers of English varied considerably in the extent to which they realized the duration contrast for vowels before voiced versus before voiceless consonants. The current study explored the perceptual consequences of this individual variability in production. Twenty native and 35 non-native listeners participated in a minimal pair identification task using stimuli from two native and two non-native talkers who differed substantially in the extent to which they realized the temporal contrast. Stimuli were consonant-vowel-consonant words that differed only in the voicing of the final consonant. Both native and non-native listeners exhibited sensitivity to variation in the production of this temporal contrast: greater identification accuracy was observed for relatively large duration differences. There was also a talker-listener interaction such that non-native listeners identified words produced by the non-native talkers more accurately than those produced by the native talkers. In a follow-up experiment, we investigated the perception-production relationship for this temporal contrast in the non-native listeners. Generally, non-native listeners who were more sensitive to this duration contrast in native-accented English showed a greater duration contrast in English production than non-native listeners who were less sensitive to this contrast. [Work supported by NIH-NIDCD.]

1aSC4. Acoustic influence in crosslanguage phone mapping: Sibilant fricative place mismatch from English to Korean. Yunju Suh (Dept. of Linguist., Stony Brook Univ., Stony Brook, NY 11794-4376, yunju.suh@sunysb.edu)

The loanword adaptation and second language production of the English voiceless palatoalveolar fricative into Korean shows an interesting mismatch in tongue position. Korean palatalizes its alveolar fricative to alveolopalatal before high front vocoids. However, English voiceless palatoalveolar fricative before a high front vowel is adapted as an alveolar

fricative with the secondary articulation of lip rounding, instead of an alveolopalatal. This paper argues that the failure to use the articulatorily closer Korean alveolopalatal fricative is due to Korean listeners' interpreting English acoustic patterns in terms of the phonetic, especially acoustic, expectations of their native language. That is, Korean listeners attend more to the peak frequency of the English fricative than to the actual tongue position, and map it to their native fricative sound with the spectral peak at the closest frequency. Data collected from female speakers of American English and Korean show that, from fricative midpoint to end, the highest intensity spectral peaks are located at similar frequencies for English palatoalveolar and Korean rounded alveolar fricatives, while those of the Korean alveolopalatal are 1500–2000 Hz higher. This serves as another piece of evidence that second language/loanword adaptation is crucially affected by fine details of L1 and L2 phonetics.

1aSC5. Listening to a non-native speaker: Adaptation and generalization. Constance M. Clarke[a] (Dept. of Psych., Univ. of Arizona, Tucson, AZ 85721)

Non-native speech can cause perceptual difficulty for the native listener, but experience can moderate this difficulty. This study explored the perceptual benefit of a brief (approximately 1 min) exposure to foreign-accented speech using a cross-modal word matching paradigm. Processing speed was tracked by recording reaction times (RTs) to visual probe words following English sentences produced by a Spanish-accented speaker. In experiment 1, RTs decreased significantly over 16 accented utterances and by the end were equal to RTs to a native voice. In experiment 2, adaptation to one Spanish-accented voice improved perceptual efficiency for a new Spanish-accented voice, indicating that abstract properties of accented speech are learned during adaptation. The control group in Experiment 2 also adapted to the accented voice during the test block, suggesting adaptation can occur within two to four sentences. The results emphasize the flexibility of the human speech processing system and the need for a mechanism to explain this adaptation in models of spoken word recognition. [Research supported by an NSF Graduate Research Fellowship and the University of Arizona Cognitive Science Program.] [a]Currently at SUNY at Buffalo, Dept. of Psych., Park Hall, Buffalo, NY 14260, cclarke2@buffalo.edu

1aSC6. Language-specific relevance of formant transitions for fricative. Anita Wagner and Mirjam Ernestus (Max-Planck-Inst. for Psycholinguist., Postbus 310, 6500 Nijmegen, The Netherlands, anita.wagner@mpi.nl)

Although the consonant inventories of Dutch, German, English, and Spanish are similar in size, the fricative inventories differ: English and Spanish distinguish labio-dental versus dental fricatives, whereas Dutch and German do not. Three phoneme-monitoring experiments investigated whether the relevance of formant transitions varies across languages, and whether it depends on the types of fricatives in a language. Native Dutch, German, English, and Spanish listeners detected a target fricative, /s/ or /f/, in nonwords. Half of the nonwords were cross spliced to produce misleading formant transitions: an /s/ replaced an /f/, or vice versa. Dutch and German listeners were unaffected by the misleading formant transitions, whereas Spanish listeners missed significantly more fricatives surrounded by misleading formant transitions; these results were obtained whether stimuli were originally spoken by a Dutch or Spanish speaker. English listeners showed the same sensitivity to formant transitions as the Spanish. Despite previous reports that formant transition cues are of negligible significance for fricative identification (Klaassen-Don, 1983), the present findings show that formant transitions are indeed relevant for listeners whose native language distinguishes labio-dental versus dental fricatives. Listeners relied on the transitions even though no dental fricatives occurred in these stimuli, and independently of the speaker's native language.

1aSC7. Pseudo-homophony in non-native listening. Anne Cutler (Max Planck Inst. for Psycholinguist., P.O. Box 310, 6500 AH Nijmegen, The Netherlands) and Takashi Otake (Dokkyo Univ., Soka, Japan)

Pseudo-homophony may result when non-native listeners cannot distinguish phonemic contrasts. Thus Dutch listeners have difficulty distinguishing the vowels of English *cattle* versus *kettle*, because this contrast is subsumed by a single Dutch vowel category; in consequence, both words may be activated whenever either is heard. A lexical decision study in English explored this phenomenon by testing for repetition priming. The materials contained among 340 items 18 pairs such as *cattle/kettle*, i.e., contrasting only in those vowels, and 18 pairs contrasting only in r/l (e.g., *right/light*). These materials, spoken by a native American English speaker, were presented to fluent non-native speakers of English, 48 Dutch Nijmegen University students, and 48 Japanese Dokkyo University students; the listeners performed lexical decision on each spoken item, and response time was measured. Dutch listeners responded significantly faster to one member of a *cattle/kettle* pair after having heard the other member earlier in the list (compared with having heard a control word), suggesting that both words had been activated whichever had been heard. Japanese listeners, however, showed no such priming for *cattle/kettle* words, but did show repetition priming across r/l pairs such as *right/light*. Non-native listeners' phonemic discrimination difficulties thus generate pseudo-homophony.

1aSC8. Issues in the measurement of perceptual assimilation. James Harnsberger and Ratree Wayland (Univ. of Florida–Gainesville, Gainesville, FL 32611, jharns@ufl.edu)

This study examined the effect of methodological variables on the fit between predicted discrimination scores based on identification data and actual discrimination data in cross-language speech perception experiments. Such variables include (1) single versus multiple talkers in discrimination test trials; (2) different discrimination test types (e.g., AX, AXB, oddity); and (3) identification tests in which stimuli are presented individually versus stimuli being presented in the same context as they appear in discrimination tests. The optimal pair of identification and discrimination tests, yielding the best match between predicted and actual discrimination scores, can be used in subsequent studies examining perceptual category structure. These methodological variables were examined by presenting American English speakers with two Hindi contrasts, [b]–[p] and breathy voiced dental-retroflex, both in initial position and in an [i], [a], or [u] context. The stimuli appeared in a range of categorial discrimination and identification tests. Early results examining the third variable listed above demonstrate that identification tests that present stimuli in the same context as they appear in corresponding discrimination test trials correlate more strongly with discrimination scores ($r=0.72$, $p<0.05$) than identification tests that present stimuli in isolation ($r=0.58$, $p<0.05$).

1aSC9. Effects of speaking rate on the perception of phonemic length contrast in Japanese. Hiroaki Kato (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, kato@atr.jp) and Keiichi Tajima (Housei Univ., Tokyo 102-8160, Japan)

Segment length is distinctive in Japanese, for example, /kaite/ (buyer) versus /kaite:/ (seabed). Such length contrasts are not necessarily categorical for non-native speakers. To study this property precisely, a series of perception experiments was conducted. A professionally trained native-Japanese speaker produced the nonsense word /erete/ at slow, normal, and fast rates with or without a carrier sentence. Either the second vowel or second consonant of each word was gradually lengthened until reaching its

longer counterpart, i.e., /ete:te/ or /eret:e/, in all rate and carrier conditions using STRAIGHT, a high-fidelity speech analysis, synthesis, and manipulation system [Kawahara *et al.*, Speech Commun. **27**, 187–207 (1999)], resulting in 12 stimulus continua. Seven native-Japanese listeners participated in a single-stimulus, two-alternative forced-choice identification task with the method of constant stimuli. The speaking rate of the presented stimuli within a session was either fixed or randomized trial by trial. Results suggest that native listeners' identification boundaries systematically altered due to changes in speaking rate, whereas their boundaries became unstable in the randomized-rate condition, especially for no-carrier stimuli. These results will be discussed from the viewpoint of second-language phoneme perception and acquisition through comparisons with results from non-native listeners. [Work supported by TAO, Japan.]

1aSC10. Spoken word recognition in English by Japanese listeners: A case of Japanese-accented and unaccented English words. Kiyoko Yoneyama (Dept. of English, Daito Bunka Univ., 1-9-1 Takashimadaira, Itabashi, Tokyo, Japan, yoneyama@ic.daito.ac.jp)

The effect of acoustic mismatch between a listener's phonological representations and speech input on spoken word recognition was examined with Japanese learners of English. Imai, Flege, and Walley (2003) recently found that native-Spanish listeners showed a larger neighborhood density effect for unaccented English words than Spanish-accented English words, whereas native-English listeners showed a larger neighborhood density effect for Spanish-accented than unaccented words. We hypothesized that phonological mismatches would occur when native-Japanese listeners respond to unaccented English words than Japanese-accented English words. Further, the effect of the mismatch would be expected to be greater for words from dense versus sparse neighborhoods because Yoneyama (2002) found that Japanese listeners showed neighborhood density effect when they listen to Japanese. This paper reports the results of the experiment where native-Japanese listeners were asked to write down English words that were presented in noise. The words differed in neighborhood density; half were Japanese-accented words that were produced by a Japanese learner of English at a beginner level, the other half were unaccented words that were produced by a native-English listener. The results replicated Imai *et al.*'s (2003) finding.

1aSC11. Effects of acoustic and semantic contexts when learning to identify L2 phonemes in words and sentences. Yuko Ikuma (ATR Human Information Sci. Lab., Kobe Univ., 2-2-2 Hikaridai Keihanna Sci. City, Kyoto 619-0288, Japan, yikuma@atr.jp) and Reiko Akahane-Yamada (Kobe Univ., Kyoto 619-0288, Japan)

Laboratory training experiment was conducted in order to examine the effect of acoustic and semantic contexts when learning second language phoneme perception. Fifty minimal pairs of English words contrasting in /r/ and /l/ were produced by native speakers of American English in three conditions; in isolation (WD), within semantically neutral carrier sentences (NS), and within semantically contextual carrier sentences (CS). Participants were native speakers of Japanese, and were divided into three groups; each was trained to identify /r/ and /l/ in one of above three conditions. In pretest, identification accuracy varied by condition in the order, NS$<$WD$<$CS, which replicated the previous study [Rothwell and Akahane-Yamada, J. Acoust. Soc. Am. **112**, 2386 (2002)]. It was also shown that the group trained with CS stimuli improved the ability to identify CS stimuli from pretest to post-test, but not WD and NS stimuli. In contrast, the effect of training using WD or NS stimuli generalized to all the stimulus conditions. These results suggest that the perception training utilizing the auditory input, in which acoustic information is the only clue to identify phonemes, is effective in cultivation of aural comprehension. Implications for foreign language education will be discussed. [Work supported by TAO, Japan.]

1aSC12. Effects of audio-visual presentation of target words in word translation training. Reiko Akahane-Yamada, Ryo Komaki, and Rieko Kubo (ATR Human Information Sci. Lab., 2-2-2, Hikaridai, Keihanna Sci. City, Kyoto 619-0288, Japan, yamada@atr.jp)

Komaki and Akahane-Yamada (Proc. ICA2004) used 2AFC translation task in vocabulary training, in which the target word is presented visually in orthographic form of one language, and the appropriate meaning in another language has to be chosen between two choices. Present paper examined the effect of audio-visual presentation of target word when native speakers of Japanese learn to translate English words into Japanese. Pairs of English words contrasted in several phonemic distinctions (e.g., /r/–/l/, /b/–/v/, etc.) were used as word materials, and presented in three conditions; visual-only (V), audio-only (A), and audio-visual (AV) presentations. Identification accuracy of those words produced by two talkers was also assessed. During pretest, the accuracy for A stimuli was lowest, implying that insufficient translation ability and listening ability interact with each other when aurally presented word has to be translated. However, there was no difference in accuracy between V and AV stimuli, suggesting that participants translate the words depending on visual information only. The effect of translation training using AV stimuli did not transfer to identification ability, showing that additional audio information during translation does not help improve speech perception. Further examination is necessary to determine the effective L2 training method. [Work supported by TAO, Japan.]

1aSC13. Temporal patterns of native Mandarin Chinese speakers' productions of English stop-vowel syllable. Yue Wang (Dept. of Linguist., Simon Fraser Univ., 8888 University Dr., Burnaby, BC V5A 1S6, Canada) and Dawn M. Behne (Norwegian Univ. of Sci. and Technol., Trondheim, Norway)

Second language (L2) production can be a kind of interlanguage, a relatively stable system bearing the nature of both the native language (L1) and L2. Within such a system sound components of a syllable may bear their own interlanguage characteristics and yet interact with the other component sounds. The present study investigates temporal patterns of L1–L2 interaction at the syllable level. Audio recordings were made of English stop-vowel syllables produced by native speakers of Mandarin who were fluent in English (ChE). Native English productions (AmE) of these syllables and native productions of Mandarin (ChM) stop-vowel syllables were acquired as native norms. Temporal measures included stop closure duration, voice-onset time (VOT), vowel duration, and syllable duration. Results show that the internal timing components of ChE often deviate from AmE, with the closure duration, VOT, and vowel duration being intermediate to AmE and ChM. However, at the syllable level, ChE productions tend to follow the overall patterns of AmE. Temporal deviations were often compensated by temporal compensation of other components in the syllable, maintaining a balanced consonant/vowel distribution. These findings have implications for a broader understanding of L2 productions.

1aSC14. Intelligibility of non-native Lombard speech for non-native listeners. Chi-Nin Li (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada)

Previous study [C-N. Li, J. Acoust. Soc. Am. **114**, 2364 (2003)] has shown that foreign-accented Lombard speech is more intelligible than normal speech when presented in noise to native English listeners. This research extends that work and examines the intelligibility of non-native English speakers' Lombard speech perceived by listeners from the same L1 background. Twelve Cantonese speakers and a comparison group of English speakers read 48 simple true and false English sentences in quiet and in 70 dB of cafeteria noise. Normal and Lombard sentences were masked with noise at a constant signal-to-noise ratio, and presented along with noise-free stimuli to eight native Cantonese speakers who assessed intelligibility by transcribing the sentences in standard English orthography. Analyses indicated that for both groups of speakers, sentences pre-

sented in noise were less well perceived than those presented without noise. The Cantonese speakers' utterances were more intelligible than were the native English productions. However, in noisy conditions, the Lombard speech of the Cantonese speakers was correctly transcribed less often than their normal utterances, and the English speakers' Lombard speech was not more intelligible than their normal speech.

1aSC15. Perceiving non-native speech: Word segmentation. Michèle Mondini and Joanne L. Miller (Psych. Dept., 125 Nightingale Hall, Northeastern Univ., Boston, MA 02115)

One important source of information listeners use to segment speech into discrete words is allophonic variation at word junctures. Previous research has shown that non-native speakers impose their native-language phonetic norms on their second language; as a consequence, non-native speech may (in some cases) exhibit altered patterns of allophonic variation at word junctures. We investigated the perceptual consequences of this for word segmentation by presenting native-English listeners with English word pairs produced either by six native-English speakers or six highly fluent, native-French speakers of English. The target word pairs had contrastive word juncture involving voiceless stop consonants (e.g., *why pink/ wipe ink; gray ties/great eyes; we cash/weak ash*). The task was to identify randomized instances of each individual target word pair (as well as control pairs) by selecting one of four possible choices (e.g., *why pink, wipe ink, why ink, wipe pink*). Overall, listeners were more accurate in identifying target word pairs produced by the native-English speakers than by the non-native English speakers. These findings suggest that one contribution to the processing cost associated with listening to non-native speech may be the presence of altered allophonic information important for word segmentation. [Work supported by NIH/NIDCD.]

1aSC16. Compensation for phonological assimilation can be triggered by nonspeech sounds. Holger Mitterer (Max-Planck-Institut für Psycholinguistik, P.O. Box 310, 6500 AH Nijmegen, The Netherlands)

Several previous experiments have shown that phonological assimilations are compensated for perceptually by a context-sensitive mechanism. For instance, the Hungarian assimilated form ''bar'' of the word ''bal'' (''left'') is recognized as such only if it occurs in a context that allows assimilation (i.e., ''balrol'' assimilated to ''barrol,'' ''from the left''), but not in other contexts (*''barnal''). This ''compensation for assimilation'' is independent of language experience: Similar results were found with Hungarian and Dutch listeners [Mitterer *et al.*, Proceedings of the 15th International Congress of Phonetic Sciences (2003)]. This suggests that an auditory mechanism underlies compensation. Nonspeech analogs of the phonetic contexts ''nal'' and ''rol'' should therefore trigger similar context effects as speech sounds: Three pairs of nonspeech analogs were played after syllables from a ''bal''-to-''bar'' continuum: First, a broadband noise was used with-analog to the trill in ''rol''—or without-analog to the nasal in ''nal''—amplitude modulation (AM) at the onset. Second, the same AM manipulation was applied to a 400-Hz tone. Third, frequency modulation replaced the AM of the 400-Hz tone. In each case, the nonspeech sounds had the same effect on identification and discrimination performance as the speech sounds.

1aSC17. The role of low-frequency hearing in speakers sensitivity to delayed auditory feedback. Dragana Barac-Cikoja and Cara Johnson (Gallaudet Univ., HMB N205F, 800 Florida Ave. NE, Washington, DC 20002)

Sensitivity to delayed auditory feedback (DAF) during speech production was assessed in six normal-hearing, fluent speakers using a two-interval, forced-choice (2IFC) adaptive procedure. Subjects repeated a syllable (either PA, TA, or KA) for approximately 2 s in two successive intervals and listened through headphones for a delay in the speech feedback. Subjects were required to identify which of the two utterances was delayed relative to the production onset. The length of the delay was changed stepwise depending on the accuracy of the subjects response. The estimates of the minimal delay yielding a 71% correct performance were obtained under three experimental conditions. The conditions varied in how the delay affected the low frequency (300 Hz) band (LFB) of the speech feedback signal: (a) LFB was delayed along with the rest of the speech spectrum; (b) LFB was separated from the rest of the spectrum by band-pass filtering, remained present in the feedback, but was never subjected to a delay; and (c) it was eliminated from the feedback by filtering and noise masking. Obtained estimates varied between 3 and 5 ms, but were not systematically related to the experimental manipulation.

1aSC18. An SII-based approach to predict the speech intelligibility in fluctuating noise for normal-hearing listeners. Koenraad S. Rhebergen and Niek J. Versfeld (Dept. of Clin. and Exp. Audiol., Academic Medical Ctr., Rm. D2-226, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands)

The speech intelligibility index (SII) is frequently used to predict the speech intelligibility for speech in a given interfering noise. However, the SII model only has been validated for speech in stationary noise. Since the SII departs from speech and noise spectra, it does not take into account any fluctuations in the masking noise. Hence, the model will yield similar SII values, regardless of the degree of fluctuation. In contrast, from the literature it is clear that normal-hearing listeners can benefit from the fluctuations in the noise. The present paper describes an SII-based approach to model speech reception thresholds (SRTs) for speech in both stationary and fluctuating noise. The basic principle of this approach is that both speech and noise signals are partitioned into small time frames. Within each time frame, the conventional SII is determined, yielding the speech information available to the listener at that time frame. Next, the SII values of these time frames are averaged, resulting in the SII for that particular condition. With the aid of SRT data from the literature, it will be shown that this approach can give a good account for most existing data.

1aSC19. Direct measurement of single and multiple passband intelligibilities: Comparison with estimates based upon the Speech Intelligibility Index. Richard M. Warren, James A. Bashford, Jr., and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, rmwarren@uwm.edu)

The intelligibility of individual bands spanning the speech spectrum is of interest for theoretical and practical reasons, and has been the subject of considerable experimental investigation. Direct measurement of passband intelligibility can be confounded with contributions from filter slopes, but by employing sufficiently high orders of FIR filtering, the present study has removed all slope contributions and measured directly intelligibilities of 1-octave and 1/3-octave passbands. Stimuli employed were based upon the same commercial recording of monosyllabic words and the same frequency bands used for the Speech Intelligibility Index (SII) [American National Standards Institute, S3.5, 1997]. SII employs an indirect procedure for estimating intelligibility: lists of band ''importance'' values are derived from intelligibility scores for high-pass and low-pass speech having incrementally varied cutoff frequencies. These importance values are correlated with intelligibility, and were transformed into intelligibility estimates using the published transfer function. Directly measured intelligibilities differ for some, but not all, SII-based intelligibility estimates for bands heard singly and in combination. Direct determination of intelligibilities of individual and multiple passbands is suggested as a simple and accurate alternative to the methods based upon SII and other indirect procedures for estimating the intelligibility of frequency-limited speech. [Work supported by NIH.]

1aSC20. Phonetic and phonological effects in production and perception of Croatian clear speech. Rajka Smiljanic and Ann R. Bradlow (Linguist., Northwestern Univ., 2016 Sheridan Rd., Evanston, IL 60208)

This paper examines the interaction of language-general, signal enhancement strategies and language-specific, phonological enhancement strategies in clear speech production and perception in Croatian, a language with a phonemic vowel length contrast and a relatively small inventory of five vowel qualities. Two native speakers of Croatian (one male, one female) read 20 nonsense sentences in conversational and clear speech. In a sentence-in-noise perception test, native Croatian listeners more accurately recognized words produced in clear than in conversational speech by both talkers. However, this clear speech intelligibility benefit was greater for the male than the female talker. Acoustic analyses showed that in clear speech both talkers enhanced the overall acoustic salience of the signal (slower, more frequent pauses, wider pitch range) and expanded the vowel space, although the male showed slightly less extensive vowel space expansion. However, only the male talker enhanced the phonemic vowel length contrast, suggesting that the observed asymmetry in the clear speech intelligibility benefit between the talkers may be due to the contribution of this phonological enhancement feature of his (but not her) clear speech production. The results suggest that speech intelligibility is most effectively enhanced by acoustic-phonetic modifications that reflect a combination of acoustic-auditory and phonological factors.

1aSC21. Vowels in clear and conversational speech: Talker differences in intelligibility for elderly hearing-impaired listeners. Sarah Hargus Ferguson (Dept. of Speech-Lang.-Hearing, Univ. of Kansas, Dole Ctr., 1000 Sunnyside Ave., Rm. 3001, Lawrence, KS 66045, safergus@ku.edu)

Several studies have shown that when a talker is instructed to speak as though talking to a hearing-impaired person, the resulting clear speech is significantly more intelligible than typical conversational speech. A recent study of 41 talkers [S. H. Ferguson and D. Kewley-Port, J. Acoust. Soc. Am. **111**, 2482 (2002)] demonstrated that for normal-hearing listeners identifying vowels in noise, the amount of improvement talkers achieve by speaking clearly varies widely. Acoustic analyses of 12 of these talkers suggested that the amount of clear speech vowel intelligibility benefit was related to specific clear speech acoustic strategies adopted by the talkers. However, an earlier paper [S. H. Ferguson and D. Kewley-Port, J. Acoust. Soc. Am. **112**, 256–271 (2002)] suggested that clear speech acoustic strategies that improve vowel intelligibility for normal-hearing listeners may not be beneficial for the group of listeners for whom clear speech is actually intended: listeners with hearing loss. The current project will explore this issue by assessing vowel intelligibility in clear and conversational speech for older hearing-impaired adults, using materials from the 12 talkers cited above. [Work supported by the University of Kansas Center for Research New Faculty General Research Fund and by NIHDCD-02229.]

1aSC22. Verbal transformation effect and the neighborhood activation model: Exploring the boundaries of the neighborhood. Peter W. Lenz, Richard M. Warren, and James A. Bashford, Jr. (Dept. of Psych., Univ. of Wisconsin—Milwaukee, P.O. Box 413, Milwaukee, WI 53201, plenz@uwm.edu)

A recorded word repeating over and over undergoes a succession of illusory changes to other words and syllables in the listener's lexicon, as well as to nonwords. This verbal transformation effect (VTE) appears to involve successive satiations of a dominant representation and serial replacements by competing representations. Early during the presentation of a VTE stimulus, reported illusory forms are typically lexical and nonlexical neighbors of the veridical stimulus (i.e., forms differing from the stimulus by a single phoneme). Interestingly, presentation of a lexical stimulus initially evokes far more reports of nonlexical neighbors while a nonlexical stimulus evokes predominantly lexical neighbors. After 6 to 7 min, the perceived forms are no longer immediate neighbors of the veridical stimulus, and differ by two phonemes on average. The neighborhood activation model (NAM) considers that identification of spoken words involves the activation of competing, phonetically similar lexical and nonlexical representations. Activation of one of these competitors reaches a critical level and that competitor enters awareness. The VTE polls the population of activated representations, providing a means of defining the functional boundaries of neighborhoods as well as the dynamics of competitor interaction. [Work supported by NIH.]

1aSC23. Frequency effects in phoneme processing. Danny Moates, Verna Stockmal, and Zinny Bond (Ohio Univ., Athens, OH 45701)

This study tested the hypothesis that phonological segments are activated during word recognition in proportion to their frequency of use, analogous to frequency effects in whole word recognition. Preliminary evidence for the hypothesis was given by Moates, Bond, and Stockmal [LabPhon **7** (2002)] using a word reconstruction task. The present study used a gating task in which progressively longer fragments of a word are presented to listeners who must identify the word after as few gates as possible. High- and low-frequency segments were contrasted by presenting them in word pairs that differed in two segments, e.g., collision–collusion, where /I/ is more frequent than /u/. We constructed 15 words pairs contrasting vowels and 16 pairs contrasting consonants (e.g., relief–release, where /s/ is more frequent than /f/). Identification judgments were gathered from 125 participants. An ANOVA showed high-frequency consonants to be identified at significantly earlier gates than their matched low-frequency consonants with both subjects and items as random factors. No such effect appeared for vowels. Also, whole words containing the high-frequency segments were not identified significantly earlier than those containing low-frequency segments. If the phoneme frequency effect is reliable, then spoken word recognition models should address it.

1aSC24. Adult perceptions of phonotactic violations in Japanese. Laurel Fais (Dept. of Psych., Univ. of British Columbia, 1401-2136 West Mall, Vancouver, BC V6T 1Z4, Canada, lfais@psych.ubc.ca), Sachiyo Kajikawa (NTT Commun. Sci. Labs., Seika-cho, Kyoto 619-0237, Japan), Janet Werker (Univ. of British Columbia, Vancouver, BC, Canada), and Shigeaki Amano (NTT Commun. Sci. Lab.)

Adult Japanese speakers ''hear'' epenthetic vowels in productions of Japanese-like words that violate the canonical CVCVCV form by containing internal consonant clusters (CVCCV) [Dupoux *et al.*, J. Exp. Psychol. **25**, 1568–1578 (1999)]. Given this finding, this research examined how Japanese adults rated the goodness of Japanese-like words produced without a vowel in the final syllable (CVC), and words produced without vowels in the penultimate and final syllables (CVCC). Furthermore, in some of these contexts, voiceless vowels may appear in fluent, casual Japanese productions, especially in the Kanto dialect, and in some, such voiceless vowels may not appear. Results indicate that both Kanto and Kinki speakers rated CVC productions for contexts in which voiceless vowels are not allowed as the worst; they rated CVC and CVCC contexts in which voiceless vowel productions are allowed as better. In these latter contexts, the CVC words, which result from the loss of one, final, vowel, are judged to be better than the CVCC words, which result from the loss of two (final and penultimate) vowels. These results mirror the relative seriousness of the phonotactic violations and indicate listeners have tacit knowledge of these regularities in their language.

1aSC25. Can Japanese listeners recognize phonemes as well as morae as word boundary cues? Takashi Otake (Dokkyo Univ., 1-1, Gakuen-cho, Soka, Saitama 340-0042, Japan, otake@dokkyo.ac.jp) and Yoko Sakamoto (Grad. School of Foreign Lang., Dokkyo Univ., Japan)

According to Shortlist (Norris *et al.*, 1995), one of the most recent universal models in spoken word recognition, metrical segmentation strategy, provides us with important cues on word boundaries, which are determined by linguistic rhythm. If this model is applied to Japanese, the cue

must be morae because it is a mora-timed language. One thing which is not fully understood yet is whether or not a phoneme boundary exists in Japanese. According to Shortlist, a possible word costraint (PWC) predicts that impossible words are inhibited, so that a phoneme boundary should not exist in Japanese. In order to test this hypothesis, two experiments were conducted using a modified version of word spotting task (search ase and hamu in nip#ase and ri#hamu) with three groups of Japanese subjects, 19 adults, 35 children, with or without Roman alphabet. The miss rates show that the subject groups could successfully find embedded words beyond both phoneme (adults:15%; children with alphabetic knowledge: 17%; children without it; 23%) and mora (adults:7%; children with alphabetic knowledge 9%; children without it; 13%) boundaries with high accuracy, suggesting that in fact Japanese could be sensitive to phoneme boundaries, but the PWC inhibits them because they are not possible words.

1aSC26. Developmental changes in the perception of syllable-final stop voicing are not explained by age-related differences in auditory sensitivity. Susan Nittrouer (Ctr. for Persons with Disabilities, Utah State Univ., 6840 Old Main Hill, Logan, UT 84322)

Two cues to voicing decisions for syllable-final stops are vocalic duration and formant transitions, especially $F1$, at syllable offset. Developmental studies show that children weight vocalic duration less and formant transitions more than adults. This study tested the hypothesis that age-related differences in auditory sensitivity can explain these developmental changes by measuring adults' and children's sensitivities to changes in the duration of nonspeech, complex tones and analogous speech stimuli. The durations of three-tone sinusoidal stimuli and natural *cob* and *cop* were varied. Duration difference limens (DLs) were similar for listeners across ages, and similar for nonspeech and speech. Also, the extent of frequency fall was varied in three-tone sinusoidal stimuli for the first tone ($T1$) only or for all three tones, as well as for $F1$ only in synthetic *buck/bug*. DLs for frequency extent were similar for listeners across ages for nonspeech stimuli when all three tones fell, but children had larger DLs than adults when only $T1$ or $F1$ fell, a finding contrary to labeling results showing that children weight this very property more than adults. Overall the evidence contradicted the hypothesis that age-related differences in auditory sensitivity can explain developmental changes in the perception of syllable-final stop voicing.

1aSC27. Implicit learning of nonadjacent phonotactic dependencies in the perception of spoken language. Conor T. McLennan and Paul A. Luce (Lang. Percept. Lab., Dept. of Psych. and Ctr. for Cognit. Sci., Univ. at Buffalo, 245 Park Hall, Buffalo, NY 14260, mclennan@buffalo.edu)

We investigated the learning of nonadjacent phonotactic dependencies in adults. Following previous research examining learning of dependencies at a grammatical level (Gomez, 2002), we manipulated the co-occurrence of nonadjacent phonological segments within a spoken syllable. Each listener was exposed to consonant-vowel-consonant nonword stimuli produced by one of two phonological grammars. Both languages contained the same adjacent dependencies between the initial consonant-vowel and final vowel-consonant sequences but differed on the co-occurrences of initial and final consonants. The number of possible types of vowels that intervened between the initial and final consonants was also manipulated. Listeners learning of nonadjacent segmental dependencies were evaluated in a speeded recognition task in which they heard (1) old nonwords on which they had been trained, (2) new nonwords generated by the grammar on which they had been trained, and (3) new nonwords generated by the grammar on which they had not been trained. The results provide evidence for listener's sensitivity to nonadjacent dependencies. However, this sensitivity is manifested as an inhibitory competition effect rather than a facilitative effect on pattern processing. [Research supported by Research Grant No. R01 DC 0265802 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.]

1aSC28. Production frequency effects in perception of phonological variation. Cynthia M. Connine and Larissa J. Ranbom (Dept. of Psych., Binghamton Univ., P.O. Box 6000, Binghamton, NY 13702, connine@binghamton.edu)

Two experiments were conducted that investigated the relationship between phonological variant occurrence frequency (based on a corpus analysis of conversational speech) and auditory word recognition. The variant investigated was an alternation between the presence of [nt] and a nasal flap (e.g., center, cen'er). The corpus analysis showed that 80% of productions are nasal flaps, with wide variability across words (from 0% for "enter" to 100% for "twenty"). In a production goodness rating experiment, listeners rated [nt] productions as better than their nasal flap counterparts. For individual items, a strong positive correlation was found between nasal flap frequency and goodness ratings: words typically produced with nasal flaps were rated as better productions. A lexical decision experiment showed that nasal flap variants were recognized more slowly and less accurately than [nt] versions. The rated quality of the nasal-flapped production was strongly correlated with the results of the lexical decision task: nasal-flapped words considered highly acceptable were recognized more quickly and accurately than words rated as poor nasal flap productions. The results demonstrate a strong relationship between experienced variant frequency and auditory word recognition and suggest that phonological variation is explicitly represented in the mental lexicon.

1aSC29. Effects of syllable onset length in determining word-likeness. Jordan Brewer, Benjamin V. Tucker, and Michael Hammond (Dept. of Linguist., Univ. of Arizona, P.O. Box 210028, 1100 E. University Blvd., Tucson, AZ 85721-0028)

Previous research on word-likeness has determined effects of neighbors, and frequency of word parts (Bailey and Hahn, 2001; Coleman and Pierrehumbert, 1997; Frisch *et al.*, 2000). Holding those contributors constant, we conducted an experiment designed to determine if there is also an independent effect of syllable onset length in determining word-likeness. Twenty students, native speakers of English, participated in a lexical decision task in which they were asked to rank tokens, e.g., /vork/, /flork/ and /strork/ (with the syllable onset varied between one and three segments) from 1 to 7. A 7 indicated a word that does not sound at all like a possible English word and 1 indicated a word that sounds like it could be a real English word (that you never heard before). A significant effect was found of syllable onset length [$F(3,20)=4.121$; $p<0.024$]. Specifically, a significant effect was found by subject but not by item between onsets of one segment versus two [$F(2,20)=6.362$; $p<0.021$], as well as one segment versus three [$F(2,20)=7.436$; $p<0.013$], but there was no effect of two versus three [$F(2,20)=0.019$].

1aSC30. Auditory phonological priming in children and adults during word repetition. Miranda Cleary and Richard G. Schwartz (City Univ. of New York Grad. Ctr., Speech and Hearing Sci., 365 Fifth Ave., New York, NY 10016, mcleary@gc.cuny.edu)

Short-term auditory phonological priming effects involve changes in the speed with which words are processed by a listener as a function of recent exposure to other similar-sounding words. Activation of phonological/lexical representations appears to persist beyond the immediate offset of a word, influencing subsequent processing. Priming effects are commonly cited as demonstrating concurrent activation of word/phonological candidates during word identification. Phonological priming is controversial, the direction of effects (facilitating versus slowing) varying with the prime-target relationship. In adults, it has repeatedly been demonstrated, however, that hearing a prime word that rhymes with the following target word (ISI=50 ms) decreases the time necessary to initiate repetition of the target, relative to when the prime and target have no phonemic overlap. Activation of phonological representations in children has not typically been studied using this paradigm, auditory-word + picture-naming tasks being used instead. The present study employed an auditory phonological priming paradigm being developed for use with

normal-hearing and hearing-impaired children. Initial results from normal-hearing adults replicate previous reports of faster naming times for targets following a rhyming prime word than for targets following a prime having no phonemes in common. Results from normal-hearing children will also be reported. [Work supported by NIH-NIDCD T32DC000039.]

1aSC31. *F*0 peaks aligned with nonprominent syllables in American English. Stefanie Shattuck-Hufnagel (RLE, MIT, 77 Massachusetts Ave., Cambridge, MA 02139, stef@speech.mit.edu), Nanette Veilleux (Simmons College), Alejna Brugos (Boston Univ.), and Robert Speer (RLE, MIT)

The occurrence of *F*0 peaks on nonprominent syllables in American English (e.g., -ing or a- in reading again) raises the question of how to label these inflection points. This pattern is not infrequent, as shown by samples from two prosodically labeled corpora of natural speech (ToBI labeled MIT Maptask and BU FM Radio News). The Maptask sample from a single speaker (235 seconds, 520 words) contained 46 H* !H* sequences; 19 had an *F*0 peak on a weak syllable between the two accented syllables. Individual speakers vary in their use of this pattern: for 3 Radio News speakers reading the same paragraph, each speaker produced 11 H* !H* sequences of which 6, 1, and 5, respectively, showed the non-accent-aligned peak. Informal listening as well as experiments described by Dilley (2004) suggest that alignment of the *F*0 peak with different nonprominent syllables between the two accents can change the perceived relative prominences of the accents, but current ToBI labels do not specify this alignment. Because *F*0 inflection points unaligned with the syllables perceived as prominent are common, and their location influences perceived prominence patterns, it is important to specify this aspect of an intonational contour. [Work supported by NIH Grants DC0075, DC02978, DC02125.]

1aSC32. The timing of speech-accompanying gestures with respect to prosody. Margaret Renwick, Stefanie Shattuck-Hufnagel, and Yelena Yasinnik (Speech Commun. Group, 36-511 MIT, 77 Massachusetts Ave., Cambridge, MA)

The question of how and whether the body movements that accompany speaking are timed with respect to the speech has often been studied, and investigators have reached different conclusions depending on the types of gestures and aspects of prosody attended to. The ToBI system for labeling pitch accents (phrase-level prosodic prominences) and intonational phrase boundaries, which provides a well-defined inventory of prosodic elements, was used to label several sound files from videotaped lectures in English. A particular type of gesture, i.e., discrete sharp rapid movements that reach a perceptually salient end point, was separately labeled for syllable location in the visual display of the same lecture samples. Preliminary analysis showed a strong correlation between this type of ''stroke-like'' gesture of the head or hands and pitch accented syllables. For example, for one speaker of Australian English, 168 of 195 strokelike gestures (86%) occurred with a pitch-accented syllable. If these observations from coarse-grained temporal labeling are confirmed by the frame-by-frame labeling now under way, it will suggest that the study of speech-accompanying gestures can provide evidence for the prosodic structure of spoken utterances, and raise the possibility that a complete model of speech production planning should include a gestural component.

1aSC33. Asking questions with focus. Fang Liu (Dept. of Linguist., Univ. of Chicago, 1010 E. 59th St., Chicago, IL 60637, liufang@uchicago.edu) and Yi Xu (Haskins Labs., New Haven, CT 06511)

This study investigates how different interrogative meanings interact with focus in determining the overall *F*0 profile of a question. We recorded eight native speakers of Mandarin producing statements, yes–no questions with and without a question particle, wh questions, incredulous questions, and confirmation questions. In each sentence, either the initial, medial, final, or no word was focused. The tonal components of the sentences are all high, all rising, all low, or all falling. *F*0 contours were extracted by measuring every complete vocal period in the initial, medial, and final disyllabic words in each sentence. Preliminary results show that in both statements and questions, the pitch range of the focused words is expanded and that of the postfocus words suppressed (compressed and lowered). However, postfocus pitch-range suppression seems less extensive in questions than in statements, and in some question types than in others. Finally, an extra *F*0 rise is often observed in the final syllable of a question unless the syllable is the question particle which has the neutral tone. This is indicative of a high or rising boundary tone associated with the interrogative meaning, which seems to be superimposed on the tone of the sentence-final syllable. [Work supported by NIDCD DC03902.]

1aSC34. Multiple effects of consonant manner of articulation and intonation type on *F*0 in English. Yi Xu and Andrew Wallace (Haskins Labs., 270 Crown St., New Haven, CT 06511, xu@haskins.yale.edu)

In this study we examine how consonant manner of articulation interacts with intonation type in shaping the *F*0 contours in English. Native speakers of American English read aloud words differing in vowel length, consonant manner of articulation and consonant position in word. They produced each word in either a statement or question carrier. *F*0 contours of their speech were extracted by measuring every complete vocal period. Preliminary results based on graphic analysis of three speakers' data suggest that there are three distinct consonantal effects: *F*0 interruption due to devoicing, a large but brief (10–40 ms) *F*0 raising at the onset of voicing, and a smaller but longer-lasting *F*0 raising throughout a large proportion of the preceding and following vowels. These effects appear to be imposed on a continuously changing *F*0 curve that is either rising-falling or falling-rising, depending on whether the carrier sentence is a statement or a question. Further analysis will test the hypothesis that these continuous curves result from local pitch targets that are assigned to individual syllables and implemented with them in synchrony regardless of their segmental composition. [Work supported by NIDCD Grant No. R01 DC03902.]

1aSC35. Acoustic correlates of perceived rhythm in spoken English. J. Devin McAuley (Dept. of Psych., Bowling Green State Univ., Bowling Green, OH 43403) and Laura C. Dilley (Harvard Univ., Cambridge, MA 02139)

Two experiments examined the relationship between speech timing and perceived rhythm for a corpus of approximately 900 spoken sentences. In experiment 1, trained listeners applied an annotation system for perceptual isochrony to the corpus. For each sentence, listeners assigned beats to syllables and judged whether the intervals between successive beats were equal or unequal, permitting the identification of perceptually isochronous speech fragments (or beat chains). For each such beat chain, the intervals between vowel onsets of successive beat syllables were determined. Overall, there was good agreement among labelers about what constituted a beat chain, with the average inter-beat-interval equal to approximately 550 ms. In experiment 2, naive listeners rated the rhythmicity of each sentence on a 6-point scale, ranging from 1 (very nonrhythmical) to 6 (very rhythmical). Sentences were judged as more rhythmic when the longest identified beat chain in experiment 1 for that sentence contained more beats and had an average IBI closer to 550 ms. IBI variability was not a significant predictor of perceived rhythmicity. The results of both

experiments will be discussed in terms of preferred tempi in perceived rhythm and entrainment models of auditory event timing (McAuley and Jones, 2003).

1aSC36. The "listener" in the modeling of speech prosody. Klaus J. Kohler (Inst. of Phonet. and Digital Speech Processing (IPDS), Univ. of Kiel, D-24098 Kiel, Germany, kjk@ipds.uni-kiel.de)

Autosegmental-metrical modeling of speech prosody is principally speaker-oriented. The production of pitch patterns, in systematic lab speech experiments as well as in spontaneous speech corpora, is analyzed in $f0$ tracings, from which sequences of H(igh) and L(ow) are abstracted. The perceptual relevance of these pitch categories in the transmission from speakers to listeners is largely not conceptualized; thus their modeling in speech communication lacks an essential component. In the metalinguistic task of labeling speech data with the annotation system ToBI, the ''listener'' plays a subordinate role as well: H and L, being suggestive of signal values, are allocated with reference to $f0$ curves and little or no concern for perceptual classification by the trained labeler. The seriousness of this theoretical gap in the modeling of speech prosody is demonstrated by experimental data concerning $f0$-peak alignment. A number of papers in JASA have dealt with this topic from the point of synchronizing $f0$ with the vocal tract time course in acoustic output. However, perceptual experiments within the Kiel intonation model show that ''early,'' ''medial'' and ''late'' peak alignments need to be defined perceptually and that in doing so microprosodic variation has to be filtered out from the surface signal.

1aSC37. Prosodic complexity and phrase length as factors in pause duration. Jelena Krivokapic (Linguist. Dept., Univ. of Southern California, 3601 Watt Way, Grace Ford Salvatori Hall 301, Los Angeles, CA 90089-1693, krivokap@usc.edu)

Research on pauses has mainly focused on predicting the likelihood of pause occurrence and on the effect of syntactic structure on pause duration within an utterance. Very little is known about what factors, apart from syntactic and discourse factors, influence the length of pauses between utterances or phrases. This experiment examines the effect of prosodic structure and phrase length on pause duration. Subjects read 24 English sentences varying along the following parameters: (a) the length in syllables of the intonational phrase preceding and following the pause and (b) the prosodic structure of the intonational phrase preceding and following the pause, specifically whether or not the intonational phrase branches into smaller phrases. In order to minimize variability due to speech rate and individual differences, speakers read sentences synchronously in dyads (Cummins, 2002; Zvonik and Cummins, 2002). The results show that length has a significant effect on pause duration both pre- and postboundary for all dyads, and that prosodic complexity has a significant postboundary effect for some dyads. The possible reasons for the observed pause duration effects and the implications of these results on the question of incrementality in speech production are discussed. [Work supported by NIH DC03172.]

1aSC38. Distinct relative $F0$ levels elicit categorical effects in $F0$ maximum and minimum alignment. Laura C. Dilley and Meredith Brown (MIT Speech Commun. Group, 50 Vassar St., 36-549, Cambridge, MA 02139, dilley@mit.edu)

A standard assumption in intonation research is that the presence and timing of fundamental frequency ($F0$) maxima and minima relative to segments are crucial phonetic characteristics of intonation patterns. The present experiment tests an alternative hypothesis that the representation of intonation patterns is based on the relative pitch levels of syllables in sequence. Synthetic stimuli were created using the phrase *Some lemonade* with an overall rising/falling or falling/rising intonation pattern. Cues to the presence and timing of $F0$ maxima and minima were eliminated by replacing the $F0$ across *lemon-* with level $F0$ contours and replacing the sonorant consonants before and after each target vowel nucleus with Gaussian noise. Four continua were created by shifting the $F0$ levels of one or both syllables in equal 0.5- or 0.75-semitone increments. Thirteen subjects imitated randomized stimuli presented over headphones. Results showed that alignment of maxima and minima in imitations was predictably related to relative $F0$ level: $F0$ maxima (minima) were aligned early in the segmental string when the first target syllable in the stimulus was higher (lower) than the following syllable; otherwise, maxima and minima were aligned late. These results provide support for models of intonation based on relative pitch levels.

1aSC39. Prosodic effects on glide-vowel sequences in three Romance languages. Ioana Chitoran (Dartmouth College Linguist., HB 6087, Hanover, NH 03755)

Glide-vowel sequences occur in many Romance languages. In some they can vary in production, ranging from diphthongal pronunciation [ja,je] to hiatus [ia,ie]. According to native speakers' impressionistic perceptions, Spanish and Romanian both exhibit this variation, but to different degrees. Spanish favors glide-vowel sequences, while Romanian favors hiatus, occasionally resulting in different pronunciations of the same items: Spanish (b[j]ela, ind[j]ana), Romanian (b[i]ela, ind[i]ana). The third language, French, has glide-vowel sequences consistently (b[j]elle). This study tests the effect of position in the word on the acoustic duration of the sequences. Shorter duration indicates diphthong production [jV], while longer duration, hiatus [iV]. Eleven speakers (4 Spanish, 4 Romanian, 3 French), were recorded. Spanish and Romanian showed a word position effect. Word-initial sequences were significantly longer than word-medial ones ($p<0.001$), consistent with native speakers more frequent description of hiatus word-initially than medially. The effect was not found in French ($p>0.05$). In the Spanish and Romanian sentences, V in the sequence bears pitch accent, but not in French. It is therefore possible that duration is sensitive not to the presence/absence of the word boundary, but to its position relative to pitch accent. The results suggest that the word position effect is crucially enhanced by pitch accent on V.

1aSC40. The difference between a question and a statement: A cross-dialect survey. Greg Kochanski, Esther Grabe, and John Coleman (Phonet. Lab., Oxford Univ., Oxford, UK)

Seven British English dialects were studied to see what prosodic distinctions are made between statements and questions in read speech. A set of Bayesian classifiers was built upon feature vectors obtained from a spectral analysis of measures of (1) f_0, (2) loudness, (3) spectral slope and (4) voicing periodicity. It was found that the prosodic information useful for the question/statement distinction is distributed broadly across the utterance, and that loudness and spectral slope can be nearly as informative as f_0 (voicing is less informative). The three important acoustic features carry somewhat less than one bit of information each, so prosodic information could be valuable to the listener, and the listener may be able to make the question/statement decision early in the utterance. The contrast differs from one acoustic property to the next: f_0 is marked primarily by slow variations. Conversely, the spectral slope and loudness measurements primarily use shorter-wavelength features, corresponding to structures that are a syllable or two long. We also find substantial differences in the prosodic information that different dialects use, and substantial differences between speakers of the same dialect. [Research supported by the UK Economic and Social Research Council, Grant RES 00-23-1049.]

Session 1aUW

Underwater Acoustics and Acoustical Oceanography: Sediment Acoustics and Scattering

Nicholas P. Chotiros, Chair
Office of Naval Research, Code 3210A, 800 North Quincy Street, Arlington, Virginia 22217-5660

Chair's Introduction—7:40

Contributed Papers

7:45

1aUW1. An experimental verification of the spherical wave effect in unconsolidated sediments. H. John Camin and Marcia Isakson (Appl. Res. Labs., Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758, misakson@arlut.utexas.edu)

Several research institutions are investigating models to properly describe high-frequency acoustic properties of unconsolidated sediments. This talk will focus on comparing experimental reflection coefficient measurements over a range of frequency (30–160 kHz) and angle to five of these theoretical sediment models: the fluid/fluid model, the visco-elastic model, the effective density fluid model (EDFM), Buckingham's microsliding model, and the Biot/Stoll poro-elastic model. Since spherical transducers were used in the experiment, spherical wave effects will be added to the models using a plane-wave decomposition method. A common set of parameters will be used in all models. Fluid and visco-elastic model parameters have been measured and verified at the test location. The Biot parameters were determined from measured values and previous inversions of plane-wave reflection data taken at the same location. Air/water interface data will be measured and computed for each model as a test case. [Work supported by ONR, Ocean Acoustics.]

8:00

1aUW2. Low-frequency sound speed reduction in unconsolidated sediments by grain rolling and sliding. Marcia Isakson and Nicholas Chotiros (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029)

In recent experiments in the Gulf of Mexico, measured low-frequency sound speeds in sandy ocean sediments were significantly less than those predicted by Woods equation, the theoretical lower limit for sound speed in unconsolidated sediment [Williams *et al.*, IEEE J. Ocean Eng. **27**(3), 413–428; Turgut and Yamamoto, J. Acoust. Soc. Am. **87**, 2376–2382 (1990)]. Therefore, the conventional treatment of sound wave velocity is apparently overlooking processes necessary to describe wave propagation in unconsolidated sediments. One possible process is a partitioning of sound wave energy into grain rolling and sliding, which leads to additional terms in the kinetic energy equation for the sand grain frame. This additional kinetic energy can be treated as an incremental virtual mass. The virtual mass terms can be evaluated given the scintillation index of the distribution of the compressional and shear stiffness variability, Poisson's ratio, and the number of contact points for a single grain. This work will present the derivation of the virtual mass terms, calculations based on likely parameter values and a comparison with the experimental measurements. [Work supported by ONR, Ocean Acoustics.]

8:15

1aUW3. Field experiments on the velocity dispersion and attenuation of acoustic waves through sandy sediments in shallow waters. Tokuo Yamamoto, Haruhiko Yamaoka (AMP, RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33176), and Junichi Sakakibara (JFE Civil Corp., Taito-ku, Tokyo, Japan)

The velocity dispersion and attenuation of acoustic waves though a sandy bottom in a shallow ocean were measured for a broad frequency band of 1–500 kHz using a buried horizontal hydrophone array and several broadband piezoelectric sources. The sands have a fairly uniform grain size of 0.22 mm and a porosity of 0.44. The experimental data were compared with the Biot theory, the BISQ theory, the squirt flow theory and the patchy-saturated theory. The data–theory comparisons indicate that the Biot theory best predicts the acoustic propagation though the sand bottom in shallow water in the medium to high frequency band of 1–500 kHz. The data–theory comparisons also confirm that the permeability, porosity, and added mass coefficient of shallow water sediments can be extracted from acoustic data. [Work supported by ONR Code 321 OA.]

8:30

1aUW4. Inversion of sediment parameters in Biot parameter space. Gopu R. Potty and James H. Miller (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett, RI 02882)

Long range sediment tomography technique [Potty *et al.*, J. Acoust. Soc. Am. **108**, 973–986 (2000)] using explosive sources has been developed and applied to data from the Primer and ASIAEX experiments. This inversion scheme estimates the compressional wave speed and attenuation matching the individual modal arrival times, i.e., inversion is carried out in the geoacoustic parameter space. This inversion scheme is now modified to invert for parameters in the sediment physical property space. Using the Biot–Stoll model as the interface between sediment physical and geoacoustic spaces, few of the most sensitive Biot parameters are inverted for. A sensitivity analysis will be carried out to identify the most sensitive Biot parameters and they will be considered as unknowns. Broad band acoustic data from wide band sources (WBS) collected on a vertical line array in frequency range 10–400 Hz will be used for the inversion. The goal of this study is to identify the sediment type (sand, silt, sandy-silt, clay, etc.) along with some physical attributes qualifying its nature (soft, hard, etc.) and the thickness of layers. This will enable us to describe the sediment characteristics over a large area with minimum number of parameters. [Work supported by ONR.]

8:45

1aUW5. Near-real-time geoacoustic characterization during the BOUNDARY 2003 experiment. Kevin D. Heaney (Lockheed-Martin ORINCON, 4350 N. Fairfax Dr., Ste. 470, Arlington, VA 22203, kevin.heaney@lmco.com) and Peter N. Nielsen (SACLANT Undersea Res. Ctr., La Spezia, Italy)

During the SACLANT Center BOUNDARY 2003 experiment on the Malta plateau, the rapid geoacoustic characterization (RGC) algorithm was applied. This algorithm uses the slope and spacing of shallow-water striations as well as the slope of the received level with range from a passing surface ship to estimate an effective geoacoustic sediment. By searching over a small set of representative sediments, using empirical relationships between grain size and compressional speed, density, and attenuation, the inversion can be carried out in near-real time. Ambient noise measurements of passing surface ships in the Malta plateau lead to geoacoustic inversion results obtained at sea that are comparable to previous work in the area where the inversions were obtained using vertical line arrays and controlled towed sources.

9:00

1aUW6. Environmental influences on the frequency dependence of effective bottom attenuation. James D. Nickila (Adv. Sonar Technol. Div. Naval Undersea Warfare Ctr., Newport, RI 02841), Kevin B. Smith (Naval Postgrad. School, Monterey, CA 93943, kbsmith@nps.navy.mil), and Gopu Potty (Univ. of Rhode Island, Narragansett, RI 02882)

Over the past several years, concern has grown over the appropriateness of bottom-attenuation models that assume a linear frequency dependence. Empirical analyses of experimental data have suggested power-law dependence with frequency exponents as high as 1.7 and above, but with large variability between geographic regions [Zhou *et al.*, J. Acoust. Soc. Am. **82**, 287–292 (1987)]. The fundamental cause of this dependence is unknown. In this analysis, the influence of the propagation and interaction with environmental variability is investigated. Specifically, a propagation model that assumes linear frequency dependence is employed which incorporates such environmental variability as range-dependent water column sound-speed profiles, bottom sound-speed gradients, bottom sound-speed and density fluctuations, and rough water/bottom interfaces. These data are then decomposed into normal-mode amplitudes at various ranges. Based on an approach similar to Potty *et al.* [J. Acoust. Soc. Am. **114**, 1874–1887 (2003)], estimates of average bottom attenuation may be inverted directly. By computing such estimates over a band of frequencies, the effective frequency dependence of the bottom attenuation can be determined as a function of range. The dominant environmental influences will also be identified and quantified.

9:15

1aUW7. The frequency dependence of the relative sensitivities of the Biot parameters. Marcia Isakson and Tracianne Neilsen (Appl. Res. Labs., Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758)

The frequency and angular dependence of the seabed reflection coefficients contain information about how the properties of the ocean bottom influence acoustic propagation. One description of the interaction between sound and sediment is the Biot–Stoll model, in which the sediment is parameterized by 13 properties. Some of these properties have a large influence on the reflection coefficients over a specific frequency band and little or no influence on reflection coefficients at other frequencies. The sensitivity of the reflection coefficients as a function of frequency to changes in the Biot–Stoll parameters are explored using a rotated coordinates inversion method. Rotated coordinates are eigenvectors obtained from an eigenvalue decomposition of the covariance matrix of gradients of the cost function to be optimized. In this case, the cost function is designed to minimize the difference between measured and modeled reflection coefficients. The resulting rotated coordinates and corresponding eigenvalues indicate the relative sensitivity of the reflection coefficients to changes in the parameters. The rotated coordinates inversion technique is applied to reflection coefficient data synthesized for a sandy bottom over subsets frequencies from 100 Hz to 1 GHz. [Work supported by ONR.]

9:30

1aUW8. Frequency dependence of modal attenuation coefficients from a simplified Biot sediment model. Wendy Saintval (Rensselaer Polytech. Inst., Troy, NY 12180, saintw@rpi.edu), William M. Carey, Allan D. Pierce (Boston Univ., Boston, MA 02215), James F. Lynch (Woods Hole Oceanogr., Inst., Woods Hole, MA 02543), and William L. Siegmann (Rensselaer Polytech. Inst., Troy, NY 12180)

Determination of the frequency behavior of modal attenuation coefficients is essential for estimation of propagation influences of poro-elastic sediments. A recent simplification of the Biot model [Pierce *et al.*, J. Acoust. Soc. Am. **114**, 2345 (2003)] provides an approach for this determination. For plane compressional waves in a homogeneous medium, the simplified theory reproduces the frequency-squared behavior of the full Biot model. However, data from a variety of shallow water locations suggests power-law exponents that are less than two. Inhomogeneities, particularly in the upper sediment layer, influence the frequency behavior. For a normally consolidated sediment (one never subjected to pressures higher than the current overburden pressure), depth profiles of porosity and sound speed can be estimated [Cederberg *et al.*, J. Acoust. Soc. Am. **97**, 2754–2766 (1995)]. We use these profiles to investigate the frequency behavior of the modal attenuations. Comparisons are provided with observations as well as previous numerical studies. [Work supported by ONR.]

9:45

1aUW9. Structure of acoustic reflection coefficient at the fluid-poroelastic interface and the medium propagating the slow wave. Keiichi Ohkawa (Div. of Appl. Marine Phys., RSMAS, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

The closed-form expressions of the acoustic reflection coefficient at the fluid-poroelastic interface exist. Due to the complexity, it is, however, rather difficult to acquire good physical insight into dependencies of the coefficient on many measurable quantities defining the fluid/porous-medium interface. Introducing effective densities allows us to obtain fine physical insight into the structure of the reflection coefficient. In the absence of the shear wave, the fundamental structure of the reflection coefficient is of the form $R=(1/Z_0-1/Z_1-1/Z_2)/(1/Z_0+1/Z_1+1/Z_2)$ where Z_0, Z_1, and Z_2 are impedances of sound, fast, and slow waves, respectively. At the normal incidence this gives the exact solution. The phase velocities of fast and slow waves, c_1 and c_2, are approximated as $c_1 \cong \sqrt{C/\rho_1}$, $c_2 \cong \sqrt{K_f/\phi\rho_2}$, where C is the Biot's coupling constant, K_f is the fluid bulk modulus, ϕ is the porosity, and ρ_1 and ρ_2 are effective densities of fast and slow waves. In air-saturated medium, ρ_2 approaches the dynamic added mass density. For the super fluid ^{4}He, c_2 becomes the same result derived by Dutta. This supports that the medium propagating the slow wave is without doubt the pore fluid in any porous media.

10:00–10:15 Break

10:15

1aUW10. Approximate single wave equation derived from Biot's porous media equations. Allan D. Pierce, William M. Carey (Boston Univ., Boston, MA 02215, adp@bu.edu), James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Mario Zampolli (NATO Undersea Res. Ctr., La Spezia, Italy)

Manipulation of Biot's equations using low-frequency and long-wavelength approximations yields wave equation with a dissipation term that predicts plane wave attenuation coefficients proportional to the square of the frequency. Examination of Biot's original derivation verifies that the wave speed in this equation is what results from Wood's theory. The derivation predicts that the fluid displacement is a predictable fraction of that of the solid matrix, and this is consistent with the behavior in Brillouin's model of an elastic spring loaded with masses of alternating magnitude. Biot's assumption that the pores must be connected is argued to be unnecessary. Unconnected pores may be elongated with constrictions, and the sound field will cause some portions of the porelets to be compressed more than other portions, so that pressure differences across constrictions force viscous flow. A heuristic derivation shows that the wave attenuation coefficient is inversely proportional to viscosity and sound speed, and directly proportional to density, the number of porelets per unit volume, the average length of porelets, and the square of the cross-sectional area of the constrictions and frequency. The heuristic theory is supported with a detailed analysis of the energy dissipation within a dumbbell-shaped porelet within an extended isotropic elastic solid.

10:30

1aUW11. Resolution cell size effects on the statistics of seafloor backscatter. Shawn F. Johnson, Anthony P. Lyons, Doug Abraham (Appl. Res. Lab., The Pennsylvania State Univ., P.O. Box 30, State College, PA 16804), and Eric Pouliquen (NATO Undersea Res. Ctr., 19138 La Spezia, Italy)

An understanding of high-frequency scattering is essential for the design and implementation of successful object detection algorithms. Due to their highly heterogeneous nature, seafloors in shallow water pose a particularly difficult problem for target detection with scattered envelope

probability density function (PDF) often having heavier tails than the Rayleigh distribution. One aspect of high-frequency scatter that has not been resolved to date, and that should also shed light on the causes of non-Rayleigh distributions, is the relation between the size of the resolution cell of the sonar system and the envelope distributions. Seafloor scattering data from transmitted CW pulses (50 and 100 kHz) and FM signals (37–65 kHz) were collected by the NATO Undersea Research Centre off of Elba Island, Italy in May 2003. Water depths were less than 100 m and bottom structures of the test sites included Posidonia Oceanica sea grass, rippled sand, and rocks. Varying the length of the CW pulses and the bandwidth of the FM signals allowed the effective resolution cell size of the sonar to be varied. The impact of changing resolution cell size on the seafloor scattering statistics will be discussed as well as implications. [Research performed under ONR Grant No. N00014-03-1-0245.]

10:45

1aUW12. Backscattering of acoustic pulses from rough surfaces: Statistical properties of first arrivals. Iosif M. Fuks (Zel Technologies, LLC and NOAA/Environ. Technol. Lab., Mail Code R/ET-0, 325 Broadway, Boulder, CO 80305) and Oleg A. Godin (Univ. of Colorado, Boulder, CO 80305)

In a number of applications, including echo sounding of the ocean bottom, nondestructive evaluation of materials, and satellite altimetry, measurements of travel time and/or shape of a backscattered pulse are employed to retrieve information about position and properties of scatterers. When the scatterer is an extensive rough surface with sufficiently large slopes, the backscattered wave consists of numerous pulses, of which the earliest arrivals are often of primary interest. In this paper, the geometrical acoustics approximation is applied to study signal intensities and the difference between the travel times of waves scattered by a random rough surface and those reflected from the mean, plane boundary. For a Gaussian statistical distribution of roughness, probability density functions (PDFs) of the travel-time differences are found for the first and second arrivals in terms of variances of the surface roughness elevation and slope. The joint PDFs of the travel times and intensities are derived. It allows one to obtain the travel-time PDF for signals exceeding a given intensity threshold. It is shown that the travel time and the intensity are strongly correlated; on average, the earlier a signal arrives, the smaller is its intensity.

11:00

1aUW13. Wedge assemblage modeling of SAX99 backscatter strengths. Richard S. Keiffer (Naval Res. Lab., Code 7181, Stennis Space Center, MS 39529-5004), Jorge C. Novarini (Planning Systems Inc., Long Beach, MS 39560-9702), Michael D. Richardson, and Kevin B. Briggs (Naval Res. Lab., Stennis Space Center, MS 39529-5004)

In SAX99 (Sediment Acoustics Experiment 1999), a number of high quality acoustic and environmental measurements were made in support of an ONR Department Research Initiative on high–frequency sound interaction in ocean sediments [M. D. Richardson *et al.* and E. I. Thorsos *et al.*, IEEE J. Ocean Eng. **26**, 4–53 (2001)]. Supporting environmental measurements include digital stereo photography of the interface roughness which were converted [A. P. Lyons *et al.*, IEEE J. Ocean Eng. **27**, 515–524 (2002)] to digital elevation maps. Using these micro-bathymetry data sets, the measured backscatter strengths from SAX99 are modeled via an extension of the wedge assemblage model [R. S. Keiffer and J. C. Novarini, J. Acoust. Soc. Am. **107**, 27–39 (2000)] that allows for density-contrast between the water and sediment but assumes no change in the sound speed. In particular, SAX99 data in which changes in the acoustic backscatter can be directly related to the time evolution of the fine-scale interface morphology will be modeled and compared. [Work supported by ONR, Program Element No. 61153N-32 and by the DoD High Performance Computing Shared Resource Center (Stennis). This document has been reviewed and approved for public release.]

11:15

1aUW14. Scattering from marine sediments with discrete inclusions. Anatoliy N. Ivakin (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105, ivakin@apl.washington.edu)

A model of scattering in sediments with discrete inclusions is developed allowing prediction of bottom reverberation, given material parameters of the sediment and inclusions and their size distribution, which can be depth dependent. The frequency-angular dependencies of the seabed scattering strength for various types of the sediment and inclusions are calculated and discussed. An environmental data set obtained at SAX99 site (near Walton Beach, FL), including the sediment particle size distribution and roughness spectra, was used to compare contributions of volume and roughness components of the seafloor scattering. It is shown, in particular, that contribution of gravel and shell inclusions and coarse sand fraction in total scattering can be dominating (over roughness) at high frequencies (about 100 kHz and higher) and grazing angles above critical (about 30 deg) while roughness at SAX99 site is likely a dominating mechanism of bottom scattering at lower frequencies and grazing angles below critical. A combined model, taking into account both roughness and volume discrete scattering, is shown to be a good descriptor of bottom reverberation in a wide frequency-angular range. Possibilities for inversion of various sediment parameters from backscattering data are discussed. [Work supported by ONR, Ocean Acoustics.]

11:30

1aUW15. A rapid method for calculating time domain scattering from a penetrable bottom. C. Feuillade (Naval Res. Lab., Stennis Space Center, MS 39529-5004)

The density contrast surface is an appropriate model for many water–seafloor interfaces, and the formalism of D. Chu [J. Acoust. Soc. Am. **86**, 1883–1896 (1989)], for the impulse response of a point source to a density contrast wedge, facilitates the extension of wedge assemblage boundary scattering models to rough penetrable surfaces. Chu's method, however, is computationally intensive, since it necessitates finding the multiple roots of a characteristic eigenvalue equation, and evaluating a corresponding power series, for each wedge apex. In this work, the direct, reflected, and diffracted field components of a density contrast wedge are considered, and the relationships between them investigated. In particular, the origin and behavior of diffractions associated with specular reflections of the source in the faces of the wedge is studied. Results show how a proper physical understanding of these phenomena leads to a simple modification of the Biot–Tolstoy theory [J. Acoust. Soc. Am. **29**, 381–391 (1957)], which then permits rapid calculations of acoustic bottom scattering and penetration in the time domain to be performed for typical sediment structures seen on the seafloor. [Work supported by ONR/NRL.]

11:45

1aUW16. *In situ* measurement of sediment acoustic properties and relationship to multibeam backscatter. Barbara J. Kraft, Luciano Fonseca, Larry A. Mayer, Glenn McGillicuddy (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03824, bjkraft@cisunix.unh.edu), Jason Ressler, Jennifer Henderson (Univ. of Rhode Island, Narragansett, RI 02882), and Peter G. Simpkin (IKB Technologies Ltd., NS B4B 1B4, Canada)

In support of the Office of Naval Research's Geoclutter Program, *in situ* acoustic and resistivity measurements were obtained using ISSAP, a device developed and built by the Center for Coastal and Ocean Mapping. The primary focus of this research is to understand the relationship between remotely measured backscatter and the acoustic properties of surficial sediments. The field area selected was Portsmouth Harbor (NH) due to the comprehensive sonar data set collected during the Shallow Water Survey 2001 conference. Seawater and surficial sediment measurements of compressional wave sound speed, attenuation, and resistivity were obtained at a large number of stations selected to represent a range of seafloor backscatter types. The ISSAP platform was configured with two orthogonal matched pairs of transducer probes operating at frequencies of 47 and 65 kHz. A Van Veen grab sampler was also used to obtain a sediment

sample at each station. Subsampling tubes were used to obtain undisturbed samples; laboratory measurements of density, compressional wave speed and attenuation, resistivity, and grain size were completed. For a small subset of samples, selected to represent a range of sediment types, measurements of permeability, shear wave speed, and attenuation were completed. [Research supported by ONR Grant No. N00014-00-1-0821.]

MONDAY AFTERNOON, 24 MAY 2004 ROYAL BALLROOM A, 2:30 TO 4:35 P.M.

Session 1pAAa

Architectural Acoustics: Multisensory Integration and the Concert Experience: How Visual Input Affects What We Hear

Jerald R. Hyde, Chair
Consultant on Acoustics, Box 55, St. Helena, California 94574

Chair's Introduction—2:30

Invited Papers

2:35

1pAAa1. Multisensory integration and the concert experience: An overview of how visual stimuli can affect what we hear. Jerald R. Hyde (P.O. Box 55, St. Helena, CA 94574)

It is clear to those who "listen" to concert halls and evaluate their degree of acoustical success that it is quite difficult to separate the acoustical response at a given seat from the multi-modal perception of the whole event. Objective concert hall data have been collected for the purpose of finding a link with their related subjective evaluation and ultimately with the architectural correlates which produce the sound field. This exercise, while important, tends to miss the point that a concert or opera event utilizes all the senses of which the sound field and visual stimuli are both major contributors to the experience. Objective acoustical factors point to visual input as being significant in the perception of "acoustical intimacy" and with the perception of loudness versus distance in large halls. This paper will review the evidence of visual input as a factor in what we "hear" and introduce concepts of perceptual constancy, distance perception, static and dynamic visual stimuli, and the general process of the psychology of the integrated experience. A survey of acousticians on their opinions about the auditory-visual aspects of the concert hall experience will be presented. [Work supported in part from the Veneklasen Research Foundation and Veneklasen Associates.]

2:55

1pAAa2. Spectro-temporal interactions in auditory-visual perception: How the eyes modulate what the ears hear. Ken W. Grant (Walter Reed Army Medical Ctr., Army Audiol. and Speech Ctr., Washington, DC 20307-5001) and Virginie Van Wassenhove (Univ. of Maryland, College Park, MD 20742)

Auditory-visual speech perception has been shown repeatedly to be both more accurate and more robust than auditory speech perception. Attempts to explain these phenomena usually treat acoustic and visual speech information (i.e., accessed via speechreading) as though they were derived from independent processes. Recent electrophysiological (EEG) studies, however, suggest that visual speech processes may play a fundamental role in modulating the way we hear. For example, both the timing and amplitude of auditory-specific event-related potentials as recorded by EEG are systematically altered when speech stimuli are presented audiovisually as opposed to auditorilly. In addition, the detection of a speech signal in noise is more readily accomplished when accompanied by video images of the speaker's production, suggesting that the influence of vision on audition occurs quite early in the perception process. But the impact of visual cues on what we ultimately hear is not limited to speech. Our perceptions of loudness, timbre, and sound source location can also be influenced by visual cues. Thus, for speech and nonspeech stimuli alike, predicting a listener's response to sound based on acoustic engineering principles alone may be misleading. Examples of acoustic-visual interactions will be presented which highlight the multisensory nature of our hearing experience.

3:15

1pAAa3. Basic research on auditory-visual interaction and listening in rooms. Frederic L. Wightman and Pavel A. Zahorik (Dept. of Psych. and Brain Sci. and Heuser Hearing Inst., Univ. of Louisville, Louisville, KY 40292)

Attending a live concert is a multisensory experience. In some cases it could be argued that hearing is the primary sense involved, but it is never the only one. Vision, smell, and even touch make important contributions to the overall experience. Moreover, the senses interact such that what one hears, for example, is influenced by what one sees, and vice-versa. This talk will address primarily the auditory aspects of the concert experience, focusing on the results of basic studies of human spatial hearing in reverberant environments, and how these results may help us understand the concert experience. The topics will include sound localization in anechoic and reverberant environments, the precedence effect, the cocktail party effect, the perception of distance, and the impact of

room acoustics on loudness perception. Also discussed will be what has been learned from empirical research on auditory-visual interactions. In this area the focus will be on the visual capture effects, the best known of which is the ventriloquism effect. Finally, the limitations of modern psychoacoustics will be addressed in connection with the problem of fully revealing the complexities of the concert experience, especially individual differences in subjective impression.

3:35

1pAAa4. Multimodal interaction in real and virtual concert halls. Pontus Larsson, Daniel Västfjäll, and Mendel Kleiner (Appl. Acoust., Chalmers Univ. of Technol., SE-412 96 Göteborg, Sweden, pontus.larsson@ta.chalmers.se)

Recently, researchers within the field of room acoustics have shown an increased interest for the understanding of how different modalities, especially vision and audition, interact in the concert hall experience. Computer auralization and virtual reality technology have brought means to efficiently study such auditory-visual interaction phenomena in concert halls. However, an important question to address is to what extent the results from such studies agree with real, unmediated situations. In this paper, we discuss some of the auditory–visual cross-modal effects discovered in previous experiments, and an account of cross-modal phenomena in room acoustic perception is proposed. Moreover, the importance of measuring simulation fidelity when performing cross-modal experiments in virtual concert halls is discussed. The conclusions are that one can expect auditory–visual interaction effects to occur in both real and virtual rooms, but that simulation fidelity might affect the results when performing experiments in virtual conditions.

3:55

1pAAa5. Comments on "intimacy" and ITDG concepts in musical performing spaces. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755, beranekleo@ieee.org).

The word "intimacy" as related to the initial-time-delay gap (ITDG) measured in halls for musical performance was born in 1961. Of two concert spaces, one was successful acoustically and the other much less so. The halls had the same number of seats and reverberation times, the principal difference being their ITDG's. To many, the hall with larger ITDG sounded arena-like and thus not "intimate." The effect of differing ITDGs appears in three of the author's books (Wiley, New York, 1962), (Acoustical Soc. of America, Melville, NY, 1996), and (Springer-Verlag, NY, 2003), with the conclusion that ITDG is an important parameter affecting the acoustical quality of concert halls and opera houses. The question is whether the word "intimacy," used in an acoustical sense, should be synonymous with ITDG. Barron (Spon, London, 1993) defines, "Intimacy refers to the degree of identification between the listener and the performance, whether the listener feels acoustically involved or detached from the music." He found from jury subjective judgments that there was little correlation between ITDG and the word "intimacy." This paper presents the author's present thinking on the usefulness of the word "intimacy" in acoustics of halls for music, and discusses experiences with the visual effect on "intimacy."

4:15

1pAAa6. Conceptual architectural/acoustical design. A. Harold Marshall (Marshall Day Acoust. Ltd., Box 5811, Wellesley, South Auckland, New Zealand)

The thinking which characterizes acoustics as a branch of physics and engineering has difficulty with the architectural design process—the process that generates a room concept in the imagination and experience of the architect. The architect has learned to "sense" the visual properties of a room as the design develops in the interaction between mind and media. Phrases such as "wanting to be" express the architectural intention but too often such intentions are dismissed as arbitrary; acoustics may then be about fixing the design with acoustical add-ons. Occasionally there is a true meeting of minds—a creative and receptive architect and an acoustician able to communicate at the level of the architectural intention. There is evidently an auditory dimension of wanting to be which is one with the visual. This paper explores the idea in several examples and concludes with suggestions for the training of acousticians.

Session 1pAAb

Architectural Acoustics and National Council of Acoustical Consultants: Student Design Competition (Poster Session)

Robert C. Coffeen, Cochair
School of Architecture and Urban Design, University of Kansas, Marvin Hall, Lawrence, Kansas 66045

Lily M. Wang, Cochair
Architectural Engineering, University of Nebraska–Lincoln, 200B Peter Kiewit Institute, 1110 South 67th Street, Omaha, Nebraska 68182-0681

Robin Glosemeyer, Chair
Jaffe Holden Acoustics, 501 Santa Monica Boulevard, Suite 606, Santa Monica, California 90401

The Technical Committee on Architectural Acoustics of the Acoustical Society of America and the National Council of Acoustical Consultants are sponsoring this Student Design Competition that will be professionally judged at this meeting.

The purpose of this design competition is to encourage students enrolled in architecture, architectural engineering, and other university curriculums that involve building design and/or acoustics to express their knowledge of architectural acoustics and building noise control in the schematic design of a buiding where acoustical considerations are of primary importance. This competition is open to undergraduate and graduate students from all nations.

The submitted designs will be displayed in this session and they will be judged by a panel of professional architects and acoustical consultants. Up to five entries will be selected for awards, one ''First Honors'' award and four ''Commendation'' awards. An award of $1,000 will be given to the entry judged ''First Honors.'' An award of $500 will be given to each of the entries judged ''Commendation.''

Session 1pAB

Animal Bioacoustics: Natural Acoustic Behavior of Animals: Session in Memory of Donald R. Griffin II

Roderick A. Suthers, Chair
Medical Sciences, Indiana University, Bloomington, Indiana 47405

Invited Papers

2:15

1pAB1. Watching bats find food: Do we classify the signals, the strategies, or the bats? James A. Simmons (Dept. of Neurosci., Brown Univ., Providence, RI 02912, james_simmons@brown.edu)

The fact that different families, genera, and even species of echolocating bats broadcast characteristic sonar signals has motivated numerous efforts to classify bats according to signal design, which has received support from correlations with both peripheral and central auditory physiology. Signal types vary according to the situations in which bats have been observed hunting for food, so this classification has been extended to the hunting strategies they use. The availability of new technical means for watching and documenting the behavior of echolocating bats in real time (thermal infrared video cameras, night-vision video with infrared illumination, video recorders with ultrasonic audio channels) makes it possible to follow individual bats long enough to observe variations in their behavior over periods of seconds to minutes. These observations reveal that at least some species nominally classified as using just one hunting strategy in fact use several strategies according to prevailing conditions, sometimes using different strategies in the course of only a few minutes. The historic inaccessibility of bats to real-time observation in the dark may have lead to exaggerated stereotyping of their behavior. [Work supported by ONR, NSF.]

2:35

1pAB2. Interaction of vestibular, echolocation, and visual modalities guiding flight by the big brown bat, *Eptesicus fuscus*. Seth S. Horowitz (Dept. of Psychiatry, State Univ. of New York at Stony Brook, HSC T-10, Rm. 086, Stony Brook, NY 11794, shorowitz@neuropop.com) and James A. Simmons (Brown Univ., Providence, RI 02912)

The big brown bat (*Eptesicus fuscus*) is an aerial-feeding insectivorous species that relies on echolocation to avoid obstacles and to detect flying insects. Spatial perception in the dark using echolocation challenges the vestibular system to function without substantial visual input for orientation. IR thermal video recordings show the complexity of bat flights in the field and suggest a highly

dynamic role for the vestibular system in orientation and flight control. Laboratory studies of flight behavior under illuminated and dark conditions in both static and rotating obstacle tests were carried out while administering heavy water (D2O) to bats to impair their vestibular inputs. *Eptesicus* carried out complex maneuvers through both fixed arrays of wires and a rotating obstacle array using both vision and echolocation, or when guided by echolocation alone. When treated with D2O in combination with lack of visual cues, bats showed considerable decrements in performance. These data indicate that big brown bats use both vision and echolocation to provide spatial registration for head position information generated by the vestibular system.

2:55

1pAB3. Bat's auditory system: Corticofugal feedback and plasticity. Nobuo Suga (Dept. of Biol., Washington Univ., One Brookings Dr., St. Louis, MO 63130)

The auditory system of the mustached bat consists of physiologically distinct subdivisions for processing different types of biosonar information. It was found that the corticofugal (descending) auditory system plays an important role in improving and adjusting auditory signal processing. Repetitive acoustic stimulation, cortical electrical stimulation or auditory fear conditioning evokes plastic changes of the central auditory system. The changes are based upon egocentric selection evoked by focused positive feedback associated with lateral inhibition. Focal electric stimulation of the auditory cortex evokes short-term changes in the auditory cortex and subcortical auditory nuclei. An increase in a cortical acetylcholine level during the electric stimulation changes the cortical changes from short-term to long-term. There are two types of plastic changes (reorganizations): centripetal best frequency shifts for expanded reorganization of a neural frequency map and centrifugal best frequency shifts for compressed reorganization of the map. Which changes occur depends on the balance between inhibition and facilitation. Expanded reorganization has been found in different sensory systems and different species of mammals, whereas compressed reorganization has been thus far found only in the auditory subsystems highly specialized for echolocation. The two types of reorganizations occur in both the frequency and time domains. [Work supported by NIDCO DC00175.]

3:15

1pAB4. Responses of inferior colliculus neurons are shaped by inhibitory projections from lower nuclei. George Pollak and Ruili Xie (Section of Neurobiology, Univ. of Texas, Austin, TX 78712, gpollak@mail.utexas.edu)

Bats are highly social animals that utilize a remarkably rich repertoire of signals for a variety of social interactions. This talk will explore how these signals are processed and represented in the inferior colliculus (IC) of Mexican free-tailed bats. Most IC neurons are selective for communication calls in that they respond only to some calls but not to others. The selectivity is due largely to inhibition. To evaluate the impact of the inhibitory projections from the dorsal or intermediate nucleus of the lateral lemniscus (DNLL and INLL), each nucleus was reversibly inactivated while responses of IC neurons to a suite of communication calls were recorded. Inactivation of both nuclei allowed these IC cells to respond to a larger number of calls, and thus become less selective than when the inhibitory innervation was intact. We therefore conclude that the inhibitory DNLL and INLL inputs to IC are functionally shaping the signal processing in the IC by suppressing some excitatory inputs. One consequence of the inhibition from these nuclei is that it shapes the selectivity of IC neurons for complex signals, allowing those IC cells to extract certain information from only some complex signals but not others. [Work supported by NIH Grant DC00268.]

3:35

1pAB5. Neural mechanisms underlying the analysis of sonar and social vocalizations: Spectral and temporal integration in the mustached bat. Jeffrey J. Wenstrup, Kiran Nataraj, Don Gans, and Kianoush Sheykholeslami (Dept. of Neurobiology, Northeastern Ohio Universities College of Medicine, 4209 State Rte. 44, Rootstown, OH 44272, jjw@neoucom.edu)

The analysis of sonar echoes by the mustached bat depends on combination-sensitive neurons that respond best when distinct spectral elements of a sonar pulse and echo occur in a particular temporal relationship. Such integrative response properties underlie direct comparisons of acoustic features in outgoing pulses and returning echoes, comparisons thought to encode pulse–echo delay and other information-bearing features of echoes. These response properties are abundant in the auditory midbrain, thalamus, and cortex of the mustached bat. Combination-sensitive neurons utilize facilitatory and inhibitory neural interactions to create selective responses to acoustic features. The different interactions (facilitatory and inhibitory) originate in different stages within the ascending auditory pathway. Inhibitory interactions arise mostly in the auditory brainstem. Facilitatory interactions arise within the inferior colliculus, but also display the results of the inhibitory interactions originating at brainstem levels. These response properties are well suited to the analysis of sonar echoes, but similar combinatorial response properties are tuned to spectral elements outside sonar frequency bands, probably to elements of social vocalizations. It is therefore likely that the neural analysis of biosonar signals shares common features with the analysis of other vocal signals. [Work supported by the National Institute on Deafness and Other Communication Disorders.]

3:55–4:10 Break

4:10

1pAB6. Pinniped bioacoustics: Atmospheric and hydrospheric signal production, reception, and function. Ronald J. Schusterman, David Kastak, Colleen Reichmuth Kastak, Marla Holt (Long Marine Lab., Univ. of California, 100 Shaffer Rd., Santa Cruz, CA 95060, rjschust@ucsc.edu), and Brandon L. Southall (Office of Protected Resources, Silver Spring, MD 20910)

There is no convincing evidence that any of the 33 pinniped species evolved acoustic specializations for echolocation. However, all species produce and localize signals amphibiously in different communicative contexts. In the setting of sexual selection, aquatic mating male phocids and walruses tend to emit underwater calls, while male otariids and phocids that breed terrestrially emit airborne calls. Signature vocalizations are widespread among pinnipeds. There is evidence that males use signature threat calls, and it is possible that vocal recognition may be used by territorial males to form categories consisting of neighbors and strangers. In terms of

mother–offspring recognition, both otariid females and their pups use acoustical cues for mutual recognition. In contrast, reunions between phocid females and their dependent pups depend mostly on pup vocalizations. In terms of signal reception, audiometric studies show that otariids are highly sensitive to aerial sounds but slightly less sensitive to underwater sounds. Conversely, except for deep-diving elephant seals, phocids are quite sensitive to acoustic signals both in air and under water. Finally, despite differences in absolute hearing sensitivity, pinnipeds have similar masked hearing capabilities in both media, supporting the notion that cochlear mechanics determine the effects of noise on hearing.

4:30

1pAB7. Cognitive processes in bird song. Jeffrey Cynx (Dept. of Psych., Vassar College, Poughkeepsie, NY 12604, chaos@vassar.edu)

Anthropomorphic hypotheses can alter previous ethological concepts. Songbirds have been traditionally categorized as open- or close-ended learners. Open-ended learners such as canaries and starlings continue to learn new songs throughout life. Close-ended learners such as song sparrows and zebra finches appear to learn song once and then repeat this song in a stereotyped or crystallized manner for the rest of their lives. Research over the last dozen years or so has produced evidence that whatever is close-ended in songbirds may be more than a little ajar. It is clear that adult song is a highly dynamic and closely monitored act. In these regards, it has a number of cognitive processes similar to human speech. Birds appear to continually monitor their own song, being able to stop in midsong if necessary. They also regulate the song amplitude given environmental and social conditions, and show song perturbations when experiencing delayed auditory feedback. However, so far as is known, close-ended learners cannot learn new song elements from a model, although there are hints to the contrary, including both behavioral and physiological results.

Contributed Papers

4:50

1pAB8. Human listening studies reveal insights into object features extracted by echolocating dolphins. Caroline M. DeLong (New College of Florida, 5700 N. Tamiami Trail, Sarasota, FL 34243, cdelong@ncf.edu), Whitlow W. L. Au (Hawaii Inst. of Marine Biol., Kailua, HI 96734), and Herbert L. Roitblat (DolphinSearch Inc., Ventura, CA 93001)

Echolocating dolphins extract object feature information from the acoustic parameters of object echoes. However, little is known about which object features are salient to dolphins or how they extract those features. To gain insight into how dolphins might be extracting feature information, human listeners were presented with echoes from objects used in a dolphin echoic-visual cross-modal matching task. Human participants performed a task similar to the one the dolphin had performed; however, echoic samples consisting of 23-echo trains were presented via headphones. The participants listened to the echoic sample and then visually selected the correct object from among three alternatives. The participants performed as well as or better than the dolphin ($M = 88.0\%$ correct), and reported using a combination of acoustic cues to extract object features (e.g., loudness, pitch, timbre). Participants frequently reported using the pattern of aural changes in the echoes across the echo train to identify the shape and structure of the objects (e.g., peaks in loudness or pitch). It is likely that dolphins also attend to the pattern of changes across echoes as objects are echolocated from different angles.

5:05

1pAB9. Biosonar performance of foraging Blainvilles beaked whales (*Mesoplodon densirostris*). Peter T. Madsen, Mark Johnson, Peter L. Tyack (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Natacha Aguilar de Soto (La Laguna Univ., Tenerife, Canary Islands, Spain), and Walter M. X. Zimmer (NATO Undersea Res. Ctr., 19138 La Spezia, Italy)

Echolocating animals like bats and toothed whales navigate and locate food by means of echoes from sounds transmitted by the animals themselves. Toothed whale echolocation has been studied intensively in captivity, but little information exists on how echolocation is used by wild animals for orientation and prey location. To expand on this issue, a noninvasive, acoustic Dtag (96-kHz sampling, 16-bit resolution) was deployed on two Blainvilles beaked whales. The tagged whales only clicked at depths below 200 m during deep foraging dives. The echolocation clicks are directional, 250-ms transients with peak energy in the 30–40-kHz band. Echoes from the seafloor and from prey items were recorded. The regular click rate is not adjusted to the decreasing echo delay from incoming prey until the target is within an approximate body length of the whale after which the click rate is increased rapidly akin to the buzz phase of echolocating bats. This suggests that the whales use different sonar strategies for operating in near versus far field modes. Changes in received echo intensities from prey targets during approaches are compared to the active gain control in the receiving system of bats and in the transmitting system of dolphins.

5:20

1pAB10. Aerial hearing sensitivity in some pinnipeds is comparable to that of humans. Colleen Reichmuth Kastak, David Kastak, Marla M. Holt, Ronald J. Schusterman (Univ. of California, Long Marine Lab., 100 Shaffer Rd., Santa Cruz, CA 95060, coll@ucsc.edu), and Brandon L. Southall (NOAA Fisheries Acoust. Prog., Silver Spring, MD 20910)

Aerial hearing sensitivity was measured in pinnipeds in a sound attenuating, hemi-anechoic chamber. Thresholds at 12 frequencies between 0.1 and 32.5 kHz were obtained behaviorally for three individuals (a California sea lion, a harbor seal, and a northern elephant seal) and compared to thresholds obtained using headphones in less controlled testing environments. The thresholds measured in the chamber revealed the expected relative changes in sensitivity with frequency; however, the absolute sensitivities were much better than had been previously measured. Harbor seal thresholds were on average 25 dB lower with best sensitivity of -2 dB (re: 20 μPa) at 3.2 kHz. Elephant seal thresholds averaged 23 dB lower with best sensitivity of 27 dB (re: 20 μPa) at 0.4 kHz. Thresholds for the California sea lion were also much lower than expected, with best sensitivity of 1 dB (re: 20 μPa) at 12 kHz. The thresholds measured for the sea lion and harbor seal rival those of human subjects at some frequencies, and suggest that previously reported aerial hearing thresholds in pinnipeds wwere significantly noise limited. Further, the results indicate that these pinnipeds have greater sensitivity in air than in water when comparisons are made in terms of sound pressure.

5:35

1pAB11. A computational model of echolocation: Transformation of spectrogram into the reflected intensity distribution for range discrimination of multiple closely spaced objects. Ikuo Matsuo, Kenji Kunugiyama, and Masafumi Yano (Res. Inst. of Elec. Commun., Tohoku Univ., 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan, matsuo@riec.tohoku.ac.jp)

Using frequency-modulated echolocation, bats can discriminate the range of objects with an accuracy of less than a millimeter. However, the echolocation mechanism is not well understood. The delay separation of three or more closely spaced objects can be determined through analysis of the echo spectrum. However, delay times cannot be properly correlated with objects using only the echo spectrum because the sequence of delay

separations cannot be determined without information on temporal changes in the interference pattern of the echoes. To illustrate this, Gaussian chirplets with a carrier frequency compatible with bat emission sweep rates were used. The delay time for object 1, $T1$, can be estimated from the echo spectrum around the onset time. The delay time for object 2 is obtained by adding $T1$ to the delay separation between objects 1 and 2 (extracted from the first appearance of interference effects). Further objects can be located in sequence by this same procedure. This model can determine delay times for multiple closely spaced objects with an accuracy of about 1 microsecond, when all the objects are located within 30 microseconds of delay separation. This accuracy is possible even with objects having different reflected intensities and in a noisy environment.

Session 1pAO

Acoustical Oceanography and Underwater Acoustics: Geoacoustic Inversion

Kyle M. Becker, Chair

Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804-0030

Contributed Papers

1:00

1pAO1. Perturbative inversion method for range-varying seabed sound speed profile estimation. Luiz L. Souza (MIT/WHOI Joint Prog. in Oceanogr./Appl. Ocean Sci. and Eng., 77 Massachusetts Ave., Rm. 5-435, Cambridge, MA 02139) and George V. Frisk (Florida Atlantic Univ., Dania Beach, FL 33004)

In a range-dependent, slowly varying medium, the acoustic field can be represented as a sum of modes that adapt to the local properties of the medium. Modal wavenumbers measured as a function of source-receiver range, as described, for example, by Souza and Frisk [J. Acoust. Soc. Am. **113**, 2204 (2003)], can be used to estimate seabed geoacoustic properties. An algorithm for estimating the compressional sound speed profile as a function of range is described. The technique consists of solving a Fredholm integral equation relating the perturbation of a background sound-speed profile to the perturbation of the modal eigenvalues [Rajan *et al.*, J. Acoust. Soc. Am. **82**, 998–1017 (1987)] in conjunction with smoothing constraints based on the stochastic regularization method. The formulation leads to a Kalman filter based algorithm, where the sound speed profile is obtained by solving the perturbative equations sequentially in range. The technique is applied to synthetic and experimental data from the modal mapping experiments (MOMAX).

1:15

1pAO2. A shallow water ocean acoustics inverse problem. David Stickler (Dept. of Mathematical Sci., New Jersey Inst. of Technol., Newark, NJ 07102)

A shallow water ocean acoustics experiment is described from which it is possible to determine in principle the scattering data necessary to recover the sound speed in the ocean bottom. This data is a "reflection coefficient." The essential assumptions are that the speed in the ocean is known but unknown in the bottom and that the measurements are made in the ocean layer.

1:30

1pAO3. Simple single path calculations for MFP tomography. Alexandra Tolstoy (ATolstoy Sci., 8610 Battailles Court, Annandale, VA 22003, atolstoy@ieee.org)

MFP tomography promises a strong computational advantage over other geoacoustic inversion methods in 3-D, variable environments. In particular, the method can simply and rapidly estimate bottom parameters in full 3-D *given* average individual path estimates of the unknown properties. The method can then perform a simple linear inversion to obtain the full volume values. But, how simple is it to estimate the individual path averages? Sometimes simple, sometimes less so. This presentation will explore this issue.

1:45

1pAO4. Tabu evaluation in geoacoustic inversion. Zoi-Heleni Michalopoulou (Dept. of Mathematical Sci., New Jersey Inst. of Technol., Newark, NJ 07102, michalop@njit.edu)

Tabu optimization has recently been proposed as a tool for geoacoustic inversion. Tabu introduces a structured use of memory in the exploration of the search space, a novel element in inversion. For performance evaluation the technique is applied to Workshop 97 test cases. These cases provide a good framework for the validation of the new approach, since results from the application of several optimization schemes to these synthetic data sets exist in the literature. Tabu is evaluated on narrow-band and broadband data in the estimation of various combinations of geoacoustic parameters. The method is successful in accurate parameter estimation and requires relatively few forward model calculations. Tabu is also evaluated in a rotated coordinate framework, in an attempt to exploit parameter correlations for faster identification of the global maximum. [Work supported by ONR.]

2:00

1pAO5. Geoacoustic inversion of M-sequence data from experiments in the South Florida Straits. Ross Chapman and Yongmin Jiang (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada)

Matched field inversion is applied to low frequency data from experiments in the South Florida Straits to estimate a sediment velocity profile. The acoustic data consist of signal envelopes derived from M-sequences that were transmitted to a vertical hydrophone array over a range of 10 km. The sound speed profile supported a strong waveguide in the deeper parts of the water column. Signal propagation for a source in the waveguide was dominated by waveguide refracted paths, and surface-reflected bottom-reflected paths that interacted eight to ten times with the bottom. The very long range, shallow water geometry presents a significant challenge for an inversion based on matched field processing. However, examination of the spatial coherence of the received signal indicated useful spatial phase information for spectral components with high signal-to-noise ratios. A Gibbs sampling approach was used to test two different geoacoustic models for the site: (a) a constant gradient sediment and (b) a

constant velocity sediment over a half space. The inversion showed a preference for slower speeds of around 1550 m/s at the sea floor, increasing to higher values up to 1700 m/s within about 100 m deeper. These values are consistent with ground truth at the site.

2:15

1pAO6. Vertical coherence of sound propagation and seabed geoacoustic inversion. Jie Yang, Xue-Zhen Zhang, Ji-Xun Zhou, and Peter H. Rogers (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332)

Array processing techniques in shallow water often require knowledge of the spatial coherence of sound propagation. However, few wideband data sets on this subject have been published. As a part of the ASIAEX program in the East China Sea, wideband sound propagation measurements were conducted in June 2001 along three radial directions and along a circle with a radius of 30 km. An earlier related paper dealt with reverberation data from this experiment [J. X. Zhou and X. Z. Zhang, J. Acoust. Soc. Am. **113**, 2204 (2003)]. This paper analyzes sound propagation vertical coherence (SPVC) as a function of frequency, range, and propagation direction. Then the SPVC is used to invert bottom sound speed and attenuation. The results show that the characteristics of the SPVC and the SPVC-inverted bottom acoustic parameters agree well with geological survey and sediment cores in the ASIAEX area. [Work supported by ONR and NNSF of China.]

2:30

1pAO7. Geoacoustic inversion of broad-band ambient noise data using undersampled and short aperture arrays. Martin Siderius, Michael Porter (SAIC, Ctr. for Ocean Res., 10260 Campus Point Dr., San Diego, CA 92121), Chris Harrison (SACLANT Undersea Res. Ctr., 19138 La Spezia (SP), Italy), and The Kauai Group (SAIC, SPAWAR, UDel, APL-UW, MPL-UCSD, SSI, UNH, USM)

Ocean ambient noise is generated in many ways such as from winds, rain and shipping. A technique has recently been developed [Harrison and Simons, J. Acoust. Soc. Am. **112** (2002)] that uses the vertical directionality of ambient noise to determine seabed properties. The ratio of beams steered towards the surface to those steered towards the bottom produces the bottom reflection loss curve. This technique was applied to data in the 200–1500-Hz band using a 16-m array. Extending this to higher frequencies allows the array length to be substantially shortened and greatly reduces interference from shipping. However, this limits the low end of the frequency spectrum since reduced aperture increases beam widths and the up/down beam ratio no longer produces reflection loss. Similarly, for high frequencies, if hydrophone spacing is greater than half-wavelength, the beamformed output is aliased and again the up/down ratio produces erroneous results. In general, frequencies much below the array design will suffer from large beams and frequencies above from undersampling. In this paper, we describe techniques for obtaining seabed properties from ambient noise measured on short or undersampled arrays. Results will be presented from the KauaiEx (July 2003) and ElbaEx (October 2003) experiments.

2:45

1pAO8. Appraisal of an energy flux model for geoacoustic inversion of ambient noise coherence data. David J. Thomson and Francine Desharnais (Defence R&D Canada—Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, david.thomson@drdc-rddc.gc.ca)

The vertical directionality of the ambient noise field in a shallow-water waveguide is highly dependent on the local geoacoustic properties of the seabed and is directly related to the vertical noise coherence observed between two closely spaced hydrophones. As a result, a direct measurement of the broadband ambient noise coherence can be used to invert for sea-bottom properties that affect propagation. In previous work, an energy flux model [D. M. F. Chapman, Proc. Inst. Acoust. **9**(4), 1–11 (1987)] for computing vertical noise coherence in shallow water was extended to include the effects of multilayered geoacoustic seabeds, refraction and absorption within the water column, sensor-pair tilt, and nondipolar radiation patterns due to near-surface but finite source depths. This modified flux-based coherence model was combined with a hybrid local/global nonlinear optimization scheme and used to estimate geoacoustic and source/sensor parameters for several synthetic data sets [F. Desharnais *et al.*, J. Acoust. Soc. Am. **113**, 2204 (2003)]. In this paper, the capability of the energy flux coherence model/geoacoustic inversion procedure is assessed for ambient noise data that were measured at several shallow-water sites over differing seabed types. Where possible, comparisons with geoacoustic properties estimated by other methods will be presented.

3:00–3:15 Break

3:15

1pAO9. Geoacoustic inversion from moving ship of opportunity in deep-water environment. David P. Knobles, Tim Scoggins, and Jack Shooter (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

This study examines the estimation of seabed geoacoustic parameters in a deep-water environment using acoustic data generated from the noise of a moving surface ship. The received sound field was recorded on a vertical line array in the North Pacific in approximately 5000 m of water. These data, which were recorded in 1973 in an experiment called Church Opal, have recently been re-digitized and made available to the underwater acoustics community. Spectrograms of time and frequency for hydrophones below the critical depth illustrate striation patterns associated with the inverse of the time interval between direct and bottom reflected rays paths. The cross-correlation of signals between phones on the VLA is thus sensitive to the structure of the seabed. A cost function of the difference between cross-phone spectra of modeled and measured received acoustic signals in the 10–600-Hz band is minimized using a simulated annealing method. Parameters that are estimated include those associated with the ship track and those associated with the seabed. Uncertainties of the various parameters are examined to determine the uniqueness of the various parameters estimated from the acoustic data. The estimated geoacoustic structure is compared to that reported in the literature. [Work supported by ONR.]

3:30

1pAO10. Iterative geoacoustic inversion. Peter Gerstoft, Chen-Fen Huang, and William S. Hodgkiss (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238)

In geoacoustic inversion we commonly invert each data block independently. But when these inversions are close in space and time it is often beneficial to use the results of the previous inversion as a starting condition for the next inversion. A natural framework for this is a Bayesian approach where the posterior inversion becomes the prior for the current inversion. It is implemented using a Metropolis–Hastings sampler. Such an approach potentially will allow us to work with lower SNR ratios, as in for example self-noise inversion. This approach has similarities to data assimilation and could lead to nowcasting of the acoustic field. [Work supported by ONR.]

3:45

1pAO11. Data error estimation for matched-field geoacoustic inversion. Michael J. Wilmut, Stan E. Dosso, and Jan Dettmer (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, mjwilmut@uvic.ca)

Nonlinear Bayesian methods have been applied to geoacoustic inversion to estimate uncertainties for seabed parameters by sampling the posterior probability density. This procedure requires quantifying the errors on the acoustic data, including both measurement and theory errors, which are generally not well known. To date, point estimates for data errors have been derived using a global maximum likelihood approach. However, this is not consistent with the Bayesian formulation, and ignores the effects of uncertainty in the error estimates and interdependencies between the data

errors and geoacoustic parameters. The Bayesian approach treats the data errors as random variables and includes them as additional parameters within the inversion. However, this increases significantly the number of unknowns and the computational effort. A third approach is to use a local maximum likelihood error estimate evaluated independently for each geoacoustic model considered in the sampling procedure. This has the benefit of not increasing the number of unknowns or computational effort, but includes some of the effects of the error uncertainties and interdependencies. The three approaches are compared for Bayesian matched-field geoacoustic inversion of both synthetic and experimental data.

4:00

1pAO12. Influence of data uncertainty on matched-field geoacoustic inversion. Chen-Fen Huang, Peter Gerstoft, and William S. Hodgkiss (Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238)

Quantifying uncertainty in matched-field geoacoustic inversion using a Bayesian approach has attracted the attention of several authors in recent years. The complete solution to the inverse problem is given by the posterior probability distribution (PPD) in the model space given good knowledge of the noise statistics. Therefore, an estimate of noise (including measurement errors and model mismatch) is essential to obtain the PPDs. In this study, several approaches to accommodating the effects of the noise are suggested, such as using different likelihood functions or treating the noise as a nuisance parameter. The posterior probability of the model parameters is sampled using a Gibbs sampler approach based on the Metropolis algorithm. Comparisons between the approaches using synthetic data and experimental data are presented. [Work supported by ONR.]

4:15

1pAO13. Bounds on errors in geoacoustic inversion: A second order analysis. William Sanders (Naval Res. Lab., Code 7185, Stennis Space Center, MS 39529)

Whereas bounding of errors in inverse problems is not generally possible, the problem has been shown to be tractable if the forward equation can be linearized about some (preferably the true) set of values and if, further, all errors are assumed Gaussian. Under these assumptions, the covariances of the *a posteriori* errors can be formulated, thus providing bounds on the uncertainty resulting from the inverse process. A previous effort analyzed the errors involved in benchmark cases and compared results to other published analyses. This surprisingly demonstrated the ability of a linearized relation to yield estimates of errors comparable to those achieved in Monte Carlo sampling. However, such a linear model neglects the interaction between model parameters. This current effort extends the analysis to second order. First and second derivatives of the acoustic field with respect to the environmental variables are calculated utilizing a parabolic equation propagation model. [Work supported by ONR.]

4:30

1pAO14. An equivalent transform method for evaluating the effect of water column mismatch on geoacoustic inversion. Ying-Tsong Lin, Chi-Fang Chen (Natl. Taiwan Univ., No. 1, Section 4, Roosevelt Rd., Taipei 106, Taiwan, yt.lin@msa.hinet.net), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

An equivalent transform method for evaluating the effect of water column mismatch on geoacoustic inversion is presented. This quantitative method is derived from perturbative inverse formulations and gives a quantitative way to evaluate how the water column mismatch error transfers into the bottom inversion results. In this paper, this method is illustrated with two test cases, in which linear and nonlinear internal waves are considered as the cause of the water column mismatch, respectively. In the first case, the range-averaged geoacoustic inversion errors due to the water column mismatch are largely eliminated after a full cycle of the linear internal wave, and the error pattern repeats with the period of the internal wave. However, in the second case, the errors are accumulated but scaled down with increasing range. Additionally, ignoring the lower-frequency components of the signal removes its corresponding contribution from the overall errors, and reduces the bottom depth influenced by water column mismatch.

4:45

1pAO15. Effects of sound speed fluctuations in the water column on geoacoustic parameter estimates. Kyle M. Becker (Appl. Res. Lab., Penn. State Univ., P.O. Box 30, State College, PA 16804-0030)

The effects of environmental variability on acoustic propagation and sonar performance in shallow water has received much attention in recent literature. Of particular interest is the influence of water column variability on acoustic propagation and the subsequent effects on geoacoustic inversion. In this paper, a modal-based inversion method is considered with the inclusion of internal waves in a shallow-water waveguide. The input data to the inversion algorithm are estimates of horizontal wavenumbers corresponding to the propagating modes of a cw point-source field. Often, wavenumber estimation is performed under the assumption that the pressure field can be represented by an adiabatic mode sum. However, it is well known that sound speed fluctuations in the water column due to internal waves promote the coupling of energy between propagating acoustic modes, thus violating the adiabatic assumption. Nevertheless, it was recently demonstrated that sound speed fluctuations in the water column due to a weak deterministic internal wave field can enhance spectral wavenumber estimates [Becker and Frisk, in *Impact of Littoral Variability on Acoustic Predictions and Sonar Performance* (Kluwer, Dordrecht, 2002)]. That work is extended to compare inversion results for wavenumber estimates obtained both with and without internal waves. [Work supported by ONR Ocean Acoustics.]

Session 1pBB

Biomedical Ultrasound/Bioresponse to Vibration: Elasticity Imaging

Elisa Konofagou, Chair
Department of Biomedical Engineering, Columbia University, 351 Engineering Terrace, New York, New York 10027

Invited Papers

1:00

1pBB1. Sonoelastography imaging: Principles and practices. Lawrence S. Taylor (BME Dept., Univ. of Rochester, 310 Hopeman Hall, Rochester, NY 14607, lstaylor@bme.rochester.edu), Deborah J. Rubens, and Kevin J. Parker (Univ. of Rochester, Rochester, NY 14607)

Vibration amplitude sonoelastography imaging is an ultrasound imaging technique in which low frequency (100–500 Hz), low amplitude shear waves are propagated deep into tissue, while real time color Doppler techniques are used to image the resulting vibration field. A radio frequency ultrasound signal is phase modulated by a vibrating particle such that the peak vibration amplitude is directly proportional to the spread (standard deviation) of its power spectrum. A key application for this technique is the detection and imaging of small lesions. Finite element studies predict that a discrete hard inhomogeneity present within a larger region of soft tissue will cause a decrease in the vibration field at the location of the inhomogeneity. The inhomogeneity is made visible as a region of low vibration in the color Doppler image. The principles and practices of this technique are reviewed and results are shown for the detection and imaging of stiff thermal lesions induced in bovine calf liver.

1:20

1pBB2. Recent advances in elastography. Jonathan Ophir (Univ. of Texas Med. School, Ultrason. Lab., 6431 Fannin, Houston, TX 77030, jophir@uth.tmc.edu)

No known modality is capable of imaging the elastic properties of tissue directly. It is therefore necessary to apply an external or internal mechanical stimulus to the tissue system, and to observe the tissue response in terms of local internal deformations. The use of ultrasound for this purpose has several important advantages that include real-time imaging, high resolution in motion estimation (1 micron), simplicity, noninvasiveness, and low cost. Elastography is performed by obtaining a set of ultrasonic echo signals from a target, subjecting the target to a small axial deformation and obtaining a second set of echo signals. Time-delay estimations along the direction of the applied load are computed. Using the gradient operator, the time-shift estimates are then converted to strain information, which is displayed in the form of a two-dimensional image named elastograms. We demonstrate that the strain distributions in tissues are well correlated with the distribution of tissue elastic moduli under certain conditions. It is also possible to obtain lateral strain elastograms and Poissons ratio elastograms, which convey additional tissue mechanical information. In this presentation we will give an overview of elastography and describe some of the newer capabilities and clinical applications. [Work supported by NIH Grant P01CA64597.]

1:40

1pBB3. Novel ultrasonic methods of imaging interstitially implanted brachytherapy seeds. Ernest J. Feleppa, S. Kaisar Alam, Jeffrey Ketterling, Andrew Kalisz, Sarayu Ramachandran, and Frederic L. Lizzi (Biomed. Eng. Labs., Riverside Res. Inst., 156 William St., New York, NY 10038)

Interstitially implanted radioactive seeds are becoming a popular and effective means of treating a wide range of diseases, primarily cancer. Brachytherapy of prostate cancer, for example, is proving to be as effective, in terms of 5-year survival, as surgery for management of gland-confined disease, and to many, the side effects of brachytherapy are less debilitating than those of surgery. While seeds are implanted under guidance from transrectal ultrasound, gland movement and distortion during implantation may lead to seed misplacement; implanted seeds are difficult to visualize ultrasonically after implantation for a variety of reasons. We are investigating a variety of ultrasonic techniques intended to overcome the limitations of conventional ultrasonic imaging, and the following three methods are showing encouraging promise: resonance/vibration imaging; modified elastography; and two-dimensional correlation. The resonance/vibration method induces natural vibrations in seeds and, using power-Doppler methods, images the movement these vibrations induce in immediately adjacent tissues. The modified-elastography method uses standard elastographic processing with higher-than-normal applied strains so that soft-tissue signals decorrelate while echo signals from rigid seeds retain correlation. The two-dimensional correlation method compares echo signals from tissue with those from a reference seed and shows seed locations by mapping the resulting correlation coefficients.

2:00

1pBB4. Acoustic radiation force impulse (ARFI) imaging: Characterizing the mechanical properties of tissues using their transient response to localized force. Kathryn R. Nightingale, Mark L. Palmeri, Amy N. Congdon, Kristin D. Frinkely (Biomed. Eng., Duke Univ., P.O. Box 90281, Durham, NC 27708-0281), and Gregg E. Trahey (Duke Univ., Durham, NC)

Acoustic radiation force impulse (ARFI) imaging utilizes brief, high energy, focused acoustic pulses to generate radiation force in tissue, and conventional diagnostic ultrasound methods to detect the resulting tissue displacements in order to image the relative mechanical properties of tissue. The magnitude and spatial extent of the applied force is dependent upon the transmit beam parameters and the tissue attenuation. Forcing volumes are on the order of 5 mm^3, pulse durations are less than 1 ms, and tissue displacements are typically several microns. Images of tissue displacement reflect local tissue stiffness, with softer tissues (e.g., fat) displacing farther than stiffer tissues (e.g., muscle). Parametric images of maximum displacement, time to peak displacement, and recovery time provide information about tissue material properties and structure. In both *in vivo* and *ex vivo* data, structures shown in matched B-mode images are in good agreement with those shown in ARFI images, with comparable resolution. Potential clinical applications under investigation include soft tissue lesion characterization, assessment of focal atherosclerosis, and imaging of thermal lesion formation during tissue ablation procedures. Results from ongoing studies will be presented. [Work supported by NIH Grant R01 EB002132-03, and the Whitaker Foundation. System support from Siemens Medical Solutions USA, Inc.]

2:20

1pBB5. Elasticity imaging of arterial wall with transcutaneous ultrasound both in longitudinal-axis and short-axis planes. Hiroshi Kanai and Hideyuki Hasegawa (Dept. of Electron. Eng., Tohoku Univ., Aramaki-Aza-Aoba 05, Sendai 980-8579, Japan)

A method for measuring regional elasticity of tissue surrounding atherosclerotic plaque is described. An ultrasonic beam was scanned with a conventional linear-type probe, and multiple layers were preset from luminal surface to adventitia of the common carotid artery (CCA) with intervals of 375 μm. By applying the method [IEEE Trans. UFFC **46**, 1229–1241 (1999)], a minute decrease of several tenths of a micrometer in thickness of each layer resulting from arrival of the pressure wave was determined. By assuming that the arterial wall is incompressible and that the blood pressure is applied normal to each layer, the elastic modulus in the circumferential direction of each layer was estimated at intervals of 75 μm in the radial direction and 150–300 μm in longitudinal direction. On the other hand, by designing the directions of ultrasonic beams so that each beam always passes through the center of the artery, the cross-sectional elasticity image in the short-axis plane was also obtained. Based on the elasticity library determined by comparing the elasticity distribution and their pathological images, each point was statistically categorized as lipid, a mixture of smooth muscle and collagen fiber, or other. By applying the method to the CCAs, soft inclusion of lipid was found for plaques.

2:40

1pBB6. Synergy of ultrasound, elasticity, and optoacoustic imaging for improved detection and differentiation of cancerous tissue. Stanislav Emelianov, Salavat Aglyamov, Jignesh Shah, Shriram Sethuraman (Dept. of Biomed. Eng., Univ. of Texas, Austin, TX 78712, emelian@mail.utexas.edu), Guy Scott, Rainer Schmitt (WinProbe Corp., North Palm Beach, FL 33408), Andrei Karpiouk, Massoud Motamedi (Univ. of Texas Medical Branch, Galveston, TX 77555), and Alexander Oraevsky (Fairway Medical Technologies, Inc., Houston, TX 77099)

The effective management of cancer requires early yet reliable detection, localization, and diagnosis. Therefore, there is a definite and urgent clinical need for an imaging technique that is widely available, is simple to perform, is safe, and that can detect and adequately diagnose cancer. In this paper we present a hybrid imaging technology based on fusion of complementary imaging modalities ultrasound, optoacoustics, and elastography to take full advantage of the many synergistic features of these modalities and thus to significantly improve sensitivity and specificity of cancer imaging. To evaluate our approach, numerical and experimental studies were performed using heterogeneous phantoms where ultrasonic, optical, and viscoelastic properties of the materials were chosen to closely mimic soft tissue. The results of this study suggest that combined ultrasound-based imaging is possible and can provide more accurate, reliable and earlier detection and diagnosis of tissue pathology. In addition, monitoring of cancer treatment and guidance of tissue biopsy are possible with combined imaging system. Practical and experimental aspects of combined imaging will be discussed with emphasis on data capture and signal/image processing algorithms. The paper will conclude with a discussion of the advantages, limitations, and potential clinical applications of the combined imaging technique.

3:00–3:30 Break

Contributed Papers

3:30

1pBB7. Studying viscoelasticity in soft tissues with supersonic shear imaging. Jeremy Bercoff, Marie Muller, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France, jeremy.bercoff@loa.espci.fr)

In this work, a 3D analytical formulation of the mechanical Green's function in a viscoelastic medium is derived and presented. Based on a Voigt model to take into account viscoelasticity, this mathematical formulation is validated experimentally using the supersonic shear imaging technique (SSI). Taking benefit of the ultrasonic remote generation of a moving shear source radiating low-frequency shear waves in the medium, this technique has been studied and validated for soft tissue elasticity mapping in previous works. It is shown here that the spatial and temporal shape of shear waves induced in soft tissues using SSI can be accurately modeled with the viscoelastic Green's function. The influences of important parameters such as viscosity, elasticity, or diffraction on the shear wave shape are carefully studied and discriminated. In a second part, taking advantage of the previous modeling, the inverse problem consisting of recovering shear elasticity and viscosity is presented and validated using the Green's function-based simulation tool. Experiments on tissue-mimicking phan-

toms presenting different viscoelastic properties are presented. The influence of out-of-plane shear propagation on the inversion algorithm is discussed.

3:45

1pBB8. A simple viscoelastic model for soft tissues in the frequency range 5–15 MHz. Xinmai Yang and Charles Church (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, University, MS 38677, xmyang@olemiss.edu)

Measurements of the shear properties of soft tissues have been performed using a simple pulse–echo reflection technique. Briefly, a shear wave, generated by a shear wave transducer, propagates through a quartz rod, is reflected at the quartz–tissue interface, backpropagates through the rod, and is received by the transducer. The phase shift and the change in the magnitude of the wave with and without a tissue sample in place are measured, and the shear mechanical impedances are calculated and the complex modulus is determined. A negative shear modulus is often obtained in this frequency range, and none of the traditional linear models can explain this result. We propose a simple extension to the Voigt model for viscoelasticity to account for the effects of mass. In this model, a negative storage modulus is possible under the conditions of high frequency and long relaxation time. When the frequency is low and relaxation time is small, the new model reduces to the classic Voigt model. Based on this model, we will present the results of measurements of shear modulus, shear viscosity, and the constant term related to mass for several soft tissues. [Work supported by the US Army.]

4:00

1pBB9. Thermal ablation monitoring using radiation force-induced steady-state tissue motion. Hesheng Wang, Yun Zhou, and Cheri Deng (Dept. of Biomed. Eng., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106, cxd54@cwru.edu)

High-intensity focused ultrasound (HIFU) increases temperature in tissue and also induces local tissue movement during HIFU ablation. Such induced motion can be exploited to detect changes in tissue stiffness to monitor HIFU treatment. To implement real-time HIFU treatment monitoring, a HIFU duration is segmented into a sequence of on and off periods. The off periods are required for unaffected interrogation and are sufficiently short to minimize interruption to the intended thermal treatment. It is demonstrated that the displacements generated during the relatively long on period reach steady state and undergo recovery during the next off period. Our study focuses on the investigation of the steady-state displacement and its recovery to detect changes in tissue elasticity without additional push pulses during an HIFU treatment exposure. An integrated model and finite difference algorithm are developed to study how the dynamic displacements are related to HIFU characteristics and tissue properties. Our experiments utilized an ultrasound imaging system. The induced displacement and its recovery are estimated from the acquired backscattered signals using a cross-correlation algorithm. Optimal HIFU on and off periods are investigated. Our results indicate that the steady-state displacements characterize the changes in tissue elasticity accompanying lesion formation during HIFU ablation.

4:15

1pBB10. Palpation strategies using the acoustic radiation force: A vectorial analysis. Jeremy Bercoff, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France, jeremy.bercoff@loa.espci.fr)

Using ultrasound focused beams, mechanical sources radiating low-frequency shear waves can be generated inside the body. Such waves propagate inside tissues at a few m/s and are sensitive to the mechanical properties of soft tissues. The ultrafast scanner developed at our laboratory (5000 images/s) is able to generate shear sources and image, in real time, the propagation of the resulting shear waves. Pushing and imaging sequences can be interleaved as desired. This great versatility enables various strategies concerning the pushing sequence. In particular, the supersonic regime, which consists of moving the shear source at a speed higher than shear wave speed, has been introduced in previous works. In this work, a detailed vectorial analysis of the supersonic regime is presented using a 3D Green's function formalism. By varying the speed and the trajectory of the shear source, conical waves of different angles, curved waves, or focused waves can be generated. Furthermore, changing the amplitude of the pushing sequences allows the temporal modulation of the resulting shear waves, leading to a complete control of the spatio-temporal shape of the shear wave. These palpation strategies are carefully studied using numerical simulations and tested on tissue-mimicking phantoms.

4:30

1pBB11. Noninvasive evaluation of complex elastic modulus of arterial vessels by ultrasound. Xiaoming Zhang, Mostafa Fatemi, and James F. Greenleaf (Dept. of Physiol. and Biomed. Eng., Mayo Clinic College of Medicine, Rochester, MN 55905)

Pulse wave velocity (PWV) is widely used for estimating the stiffness of an artery. It is well known that a stiffened artery can be associated with disease and aging. Usually, PWV is measured using the foot-to-foot time-delay method. However, the foot of the pulse wave cannot be accurately determined due to reflected waves. In addition, PWV is an average indicator of artery stiffness between the two measuring points. We propose to generate a wave in the artery wall using localized ultrasound radiation force. The wave velocity can be measured accurately by the phase change versus distance over a few millimeters. A mathematical model is developed by which the real part of the complex Young's modulus of the artery can be determined from measured wave velocities. The imaginary part of the complex modulus can be assessed by the wave amplitude decay over distance. Experiments were conducted on a pig carotid artery in gelatin. The measured wave velocity is about 3 m/s at 100 Hz and 6.5 m/s at 500 Hz. The real part of complex modulus is estimated to be 300 kPa. The present technique offers a new method for estimating the complex elastic modulus of the artery over a short length.

4:45

1pBB12. Improvement in elastographic signal-to-noise ratio and resolution using coded excitation in elastography. Remi Souchon, Jean-Christophe Bera, Agnes Pousse, and Jean-Yves Chapelon (INSERM U556, 151 cours Albert Thomas, 69424 Lyon cedex 03, France, souchon@lyon.inserm.fr)

Coded excitation has been used in conventional sonography to improve the sonographic signal-to-noise ratio independently of resolution. In elastography, a tradeoff exists between spatial resolution and the elastographic signal-to-noise ratio (SNRe). In the present work, the use of coded excitation was investigated to remove this ambiguity in elastography. Both numerical simulations and phantom experiments were carried out to estimate the SNRe in a homogeneous material and the resolution in a material containing calibrated inclusions. These results were compared with those obtained using conventional pulse excitation. Various codes (Golay, Barker, chirp) and code lengths were tested. The numerical simulations used a simple 1D backscattering model to show the theoretical effects of coded excitation. Experiments were carried out using a more realistic setup based on a sector-scan imaging probe. In the absence of sonographic noise, simulations showed that codes induced only a slight decrease in SNRe at no cost in resolution. When sonographic noise was added into the model, a large improvement in SNRe was obtained at constant resolution. The experimental results corroborated these findings. [Work supported in part by National Cancer Institute (USA) Program Project Grant P01-CA64597.]

5:00

1pBB13. Strain imaging with intravascular ultrasound: An *in vivo* study. Christian Perrey, Helmut Ermert (Lehrstuhl fuer Hochfrequenztechnik, 44780 Bochum, Germany, christian.perrey@rub.de), Waldemar Bojara, Stephan Holt, and Michael Lindstaedt (Bergmannsheil, 44789 Bochum, Germany)

The evaluation of mechanical properties of coronary plaques is of high interest for the assessment of coronary diseases. Intravascular ultrasound (IVUS) can be used to visualize strain in coronary tissue. In this study, strain imaging is performed using an IVUS system with a 40-MHz rotating single-element transducer. Radio frequency (rf) data are acquired during *in vivo* examinations and sampled at 100 MHz. Image frames are stored consecutively during 3 s at a frame rate of 30/s. Data are recorded at different levels of tissue compression. The required pressure difference is caused by natural pulsatile blood flow. The strain imaging algorithm estimates radial strain from rf data based on frame-to-frame correlation. Rotating transducers often show nonuniform rotational distortion (NURD), which leads to misaligned echo lines in consecutive frames. This results in lateral motion artifacts and causes decorrelation. This effect is reduced by lateral motion correction based on block-matching algorithms. Results show that strain imaging can successfully be performed *in vivo* with data acquired predominantly in diastole. Different coronary tissue regions can be identified by local strain variations. If NURD is present, strain image quality is degraded. In some cases NURD is reduced by repositioning the transducer.

5:15

1pBB14. Automatic error correction in quantitative elasticity imaging. Paul E. Barbone (Dept. of Aerosp. & Mech. Eng., Boston Univ., Boston, MA 02215, barbone@bu.edu)

We consider the inverse problem associated with imaging the elastic shear modulus of soft tissue. We work from the premise that tissue deformation resulting from external forces has been imaged by ultrasound. From this measured deformation, an assumed form of the tissue's constitutive equation (e.g., linear elastic), and the law of conservation of momentum we arrive at the inverse problem. We present mathematical results that show how many data are required to reliably reconstruct the shear modulus. Further, we present a consistency condition that must be satisfied by the measured deformation field, provided that the tissue obeys the laws of linear elasticity. From this consistency condition, we show how to identify the existence of errors in data. These can then be corrected, under the assumption that the measured deformation field is close to the actual field. We present several computed examples. [Work supported by NSF via CenSSIS ERC.]

MONDAY AFTERNOON, 24 MAY 2004 LIBERTY 1/2, 1:25 TO 3:00 P.M.

Session 1pNS

Noise: Noise Source Characterization

Richard H. Lyon, Chair
RH Lyon Corp, 691 Concord Avenue, Cambridge, Massachusetts 02138

Chair's Introduction—1:25

Contributed Papers

1:30

1pNS1. The noise-cooling tradeoff in electronic equipment. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

The noise produced by cooling air passing through electronics packages arises from two sources. One source is the noise of the air-moving fan of either an axial or centrifugal type. This noise may have tonal components and both those components and the broad band noise are dependent on the way that the fan is placed in the unit and on how close the operation is to the design operating point. Often, this can be the dominant noise source. The other source produces random noise due to the turbulent flow of air through the unit. Because the turbulent airflow is also responsible for heat transfer between the components and the air stream, we can regard this part of the noise as the irreducible noise due to cooling. If fan noise were eliminated, this part of the noise must remain. There is a relation therefore between the irreducible noise and the cooling of the unit. But the fan noise must also be considered. The relation between total airflow related noise and cooling requirements is developed in this paper.

1:45

1pNS2. Design and application of a perforator element muffler for reducing noise from fan-trays in telecommunication enclosures. Hugh Holness and M. G. Prasad (Dept. of Mech. Eng., Stevens Inst. of Technol., Hoboken, NJ 07030, hholness@stevens.edu)

Generally the mufflers for reducing noise from fan-trays in telecommunication enclosures use sound absorptive materials in their design. Our objective in this work is to design a perforator element muffler that will produce similar acoustical performance without increasing pressure drop in the system as compared to a generally used muffler with sound absorptive material. The work includes parametric studies on the influence of perforated element on noise attenuation and pressure drop. The advantages and disadvantages of reactive perforator element mufflers over the dissipative mufflers in such applications will be studied. The studies include both design and experimental work.

2:00

1pNS3. Road traffic noise impact assessment based on the difference between existing and future traffic noise levels. Chetlur G. Balachandran, Sandor Juhasz, Arthur Morrone (Parsons Brinckerhoff, One Penn Plaza, New York, NY 10117, Balachandr@pbworld.com), and Noemi Castillo (HDR Inc., White Plains, NY 10004)

New York City has established standards for impact assessment based on the difference in traffic noise levels between the existing and future conditions. This difference is directly related to the logarithmic ratio of passenger car equivalents under the two conditions. The FHWA Traffic Noise Model 2.0 predicts noise levels from data based on the number of vehicles classified by category, average vehicle speed, and the distance to receptor. The model results can be used to determine the change in noise level between the existing and future conditions. Both approaches give essentially the same result. In heavily populated urban centers like New York City, where the measured baseline existing noise is not only from road traffic but also from many other noise sources such as air-

conditioning units, people noise, construction activities, etc., a new approach is needed to determine the change between the existing and future noise levels. This approach utilizes data obtained from field noise measurements which are conducted with and without trucks. This realistic procedure can be used to determine the number of new heavy trucks that could be introduced into the existing traffic stream without exceeding impact thresholds set by the New York City Noise Regulations.

2:15

1pNS4. Determination of sound power levels of some pyrotechnic devices using sound pressure measurements. Weixiong Wu (117 E. 29th St., New York, NY 10016)

A noise measurement study was conducted at five reservoirs in New York State during the winter and summer of 2003 to determine the sound power levels generated by some specific pyrotechnic devices (bangers, screamers, and CAPAs) using sound pressure techniques. The study was performed in support of an environmental impact statement (EIS) that defined the areas around each of these reservoirs where any significant impacts would occur as a result of the pyrotechnic operations. Sound characteristic data for each pyrotechnic device was collected at short distances with a direct line of sight to the pyrotechnic sources. The sound pressure levels for each pyrotechnic device were measured in two conditions (winter and summer) to examine the accuracy of measured data. The sound power levels, including both A-weighted and C-weighted 1/1 octave band values, were calculated based upon the measured sound pressure levels. With the absence of literature and manufacturing data supports, the values of this study were believed to be the most comprehensive emission data for the EIS noise analysis.

2:30

1pNS5. Environmental noise impact of modern wind farms. Andrew Piacsek and Greg Wagner (Central Washington Univ., 400 E. University Way, Ellensburg, WA 98926-7422, piacsek@cwu.edu)

Electric power production from wind turbines has increased substantially during the past few years due to the growing emphasis on renewable energy sources and more efficient wind turbine technology. Although modern turbines are significantly quieter than early models, wind farms that are proposed near residential areas generate concern about potential noise issues. The present study consists of two parts: (1) the measurement of sound levels within a 2-km radius of the existing Nine Canyons wind farm near Richland, WA, and (2) the application of an outdoor sound propagation model to predict noise levels, both at Nine Canyons and in the vicinity of a proposed farm near Ellensburg, WA. At most locations within the Nine Canyon site, recorded sound levels were less than predicted levels, with the exception of some downwind sites that were lower in elevation than the source. Noise levels were greatest downwind from the turbines, but never exceeded 50 dBA beyond 500 m from the nearest turbine. In many cases, wind noise at the microphone exceeded noise levels from the turbines.

2:45

1pNS6. Acoustic project for installation of motor generator group by means of computer simulation. Jose C. Ferreira and Paulo T. Zannin (Lab Acustica Ambiental, Universidade Federal do Parana, Dept. Engenharia Mecanica, Centro Politecnico)

This work presents an acoustical project for the installation of a motor generator group of electricity in a hotel by means of computer modeling. The noise levels at the site have been obtained without the motor generator group, and via the computer modeling it has been deduced how these levels would be after the installation of the equipment. A possible solution to mitigate the noise impact the equipment would cause on the neighborhood has been indicated, and it has been predicted how the impact would be reduced after the implantation of this solution.

MONDAY AFTERNOON, 24 MAY 2004 NEW YORK BALLROOM B, 1:00 TO 4:45 P.M.

Session 1pPAa

Physical Acoustics: Recent Advances in Buried Landmine Detection II

James M. Sabatier, Cochair
University of Mississippi, National Center for Physics Acoustics, Coliseum Drive, University, Mississippi 38677

Bradley Libbey, Cochair
Night Vision and Electronic Sensors Directorate, 10221 Burbeck Road, Fort Belvoir, Virginia 22060

Invited Papers

1:00

1pPAa1. Recent developments in laser Doppler velocimeter for landmine detection. Amit Lal, Vyacheslav Aranchuk, and Cecil Hess (MetroLaser, 2572 White Rd., Irvine, CA 92614)

This presentation discusses laser Doppler vibrometer (LDV) instruments developed by MetroLaser that are specifically tailored to work with the acoustic/seismic technique for buried landmine detection and data obtained with these systems under various field conditions. The first system is a simple and rugged diode laser LDV that measures a single point on the ground. Multiple diode systems or a scanning mirror placed in front of the diode LDV are some of the strategies used to measure multiple points. The second system is a multi-beam LDV that uses a single YAG laser and diffraction optical elements to create an array of laser beams and thus multiple measurement points on the ground. The data obtained with LDV consist of velocity profiles at each point of the surface illuminated by each of the beams. By interrogating multiple points (with multiple vibrometers or with a multi-beam vibrometer) the system can ascertain with high probability the presence of a buried landmine. The systems exhibit a velocity resolution of about 1 micron per second, which is suitable for landmine detection. Low-frequency (100 kHz) carrier is employed in the multi-beam system to further enhance its performance with the acoustic/seismic technique.

1:20

1pPAa2. Investigation of a combined seismic, radar, and induction sensor for landmine detection. Waymond Scott, Jr., Kangwook Kim (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332-0250, waymond.scott@ece.gatech.edu), and Gregg Larson (Georgia Inst. of Technol., Atlanta, GA 30332-0405)

An experimental system to collect co-located ground penetrating radar (GPR), electromagnetic induction (EMI), and seismic data was developed to investigate the benefits of the fusion of these sensors. In the experiments, a range of mines and clutter objects were buried at various depths in the sandbox at Georgia Tech. Multiple burial scenarios were investigated with a variety of antipersonnel and antitank mines and typical clutter objects. The seismic system used in these experiments is an extension of our existing seismic mine detection system. The system uses electrodynamic shakers to generate a seismic wave which propagates across the simulated minefield, and a specially designed radar is used to measure the displacement of the surface caused by the seismic wave. The GPR makes use of modified resistive-vee antennas and operates over the frequency range of 500 MHz to 8 GHz. These antennas are very clean in that they have very little self-clutter and very low radar cross section to lessen the reflections between the ground and the antennas. The EMI sensor collects broadband data (300 Hz to 100 kHz) so that the relaxation frequencies of the buried targets can be used to aid discrimination. [Work supported by ARO.]

1:40

1pPAa3. Ground-penetrating radar for landmine detection applications. Marshall Bradley and Michael Duncan (Planning Systems, Inc., 115 Christian Ln., Slidell, LA 70458, mduncan@mail.psislidell.com)

The use of ground-penetrating radar (GPR) for landmine detection and its relationship to acoustic–seismic (A/S) techniques is discussed. While GPR is very effective against buried metal targets, modern plastic-cased landmines are much more difficult for GPR to detect. This difficulty is due to the low dielectric contrast between the landmine material and the soil itself. Recent advances in radar technology and signal-processing algorithms have improved GPR performance against buried plastic landmines. Several multichannel and synthetic aperture GPR systems have been developed that are capable of acquiring broad spectrum data over densely sampled spatial grids. Volumetric images of the subsurface produced by these systems are processed using automated-target-recognition (ATR) routines in order to determine the presence/absence of a landmine target. While GPR systems are capable of performing well on their own, the fusion of GPR data with data from an additional ''orthogonal'' sensor greatly increases the overall probability of detection (Pd) and reduces the false-alarm rate (FAR). Acoustic–seismic (A/S) sensors provide this orthogonality by exploiting mechanical properties of the target as opposed to the electromagnetic properties detected by GPR.

2:00

1pPAa4. Laser Doppler vibrometer-based acoustic/seismic technique and GPSAR technique for co-located landmine detection. Ning Xiang (Dept. of Elec., Computer, and Systems Eng., and School of Architecture, Rensselaer Polytechnic Inst., Troy, NY 12180) and Marshall Bradley (Planning Systems, Inc., Slidell, LA 70458)

Recent success in using a laser Doppler vibrometer (LDV) based acoustic-to-seismic landmine detection [J. M. Sabatier and N. Xiang, IEEE Trans. Geosci. Remote Sensing **39**, 1146–1154 (2001); N. Xiang and J. M. Sabatier, J. Acoust. Soc. Am. **113**, 1333–1341 (2003)] and a ground penetrating synthetic aperture radar (GPSAR) suggested a novel configuration of fused sensors comprised of a LDV-based A/S detection sensor and a GPSAR. Extensive field experiments revealed that these two technologies can be considered orthogonal. When used in concert, a fused configuration may significantly improve the probability of detection and reduce the false alarm rate. They function best against different types of landmines under different burial conditions because they exploit disparate phenomena to detect mines. In order to optimize the fused detection ability using the two modalities, co-located field experiments have been conducted using both LDV-based A/S sensor and a GPSAR sensor. This paper will discuss the comparative experimental study using the recent co-located field scanning results.

2:20

1pPAa5. Landmine detection from acoustic/seismic signals using ordered weight-averaging operators. Paul Gader, Roopnath Grandhi, Ali Hocaoglu (301 CSE Bldg, CISE Dept., Univ. of Florida, Gainesville, FL 32611, pgader@cise.ufl.edu), James Keller, and Tsaipei Wang (Univ. of Missouri–Columbia, Columbia, MO 65211)

Acoustic/seismic systems demonstrate significant potential for reliable detection of landmines, particularly when coupled with ground-penetrating radar (GPR). Acoustic/seismic systems provide complementary information to GPR. One difficulty in developing detection algorithms for acoustic/seismic systems is that the frequency bands at which information occurs vary according to factors such as mine depth and mechanical properties of the soil, which are not known during mine detection. Ordered weighted-averaging (OWA) systems offer the potential for robust detection of landmines under different conditions because they naturally provide the capability to process the best collection of frequency bands in a fashion that can be optimized. In this paper, a processing framework is presented that combines OWA operators for feature analysis with decision making. This framework can be optimized using acoustic/seismic signals as well as acoustic/seismic signals combined with GPR. In the latter case, the optimization can help to provide the appropriate mechanism for combining the complementary information from both sensors. Results are presented using real data from both mines and other buried objects collected at an outdoor test site.

2:40

1pPAa6. An acoustic landmine detection confirmatory sensor using continuously scanning multibeam LDV. James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, sabatier@olemiss.edu), and Vyacheslav Aranchuk (MetroLaser, Inc., Irvine, CA 92614-6236)

Acoustic-to-seismic coupling technology using an LDV as a vibration sensor has proved itself as a potential confirmative sensor for buried landmine detection. One of the most important objectives of this technology is to increase the speed of measurements. Traditionally used point-by-point scanning LDVs cannot provide fast measurements. A moving platform that uses 16 LDVs and a continuously scanning single-beam LDV has been used to increase the speed of detection of buried landmines. Recently a multibeam LDV simultaneously probing 16 positions on the ground has been developed and successfully used for landmine detection. In this work, we report the use of a continuously scanning multibeam LDV as a confirmatory sensor for acoustic landmine detection. The multibeam LDV simultaneously illuminates the ground in 16 points spread over a 1-m line. A scanning mirror moves all 16 laser beams across the line. An airborne sound source in the frequency range of 80–300 Hz has been used to excite vibrations in the ground. The system enables scanning a 1-m by 1-m area and provides the vibrational pattern at the surface of the ground in a much shorter time than with previous scanning techniques.

2:55

1pPAa7. A different approach for processing multibeam laser Doppler vibrometer (LDV) data. Ronald A. Wagstaff and Kenneth E. Gilbert (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, rwagstaf@olemiss.edu)

Exciting the ground with an acoustic tonal projected by a loudspeaker is a well-known method for detecting buried landmines. The subsequent ground motion is measured with a laser Doppler vibrometer (LDV). The LDV data contain the tonal in a frequency-modulated form. One approach for demodulating the data and extracting the tonal uses a Hilbert transform. The ground velocity can be obtained from these data to identify mine presence or absence. An alternate approach to mine detection is to perform consecutive fast Fourier transforms on the modulated LDV data, and to average the output powers in each spectral bin. This results in a ground velocity distribution function in the spectrum. The proximity of the beams to a mine (over, near, not near) can be determined from the width of the velocity distribution functions. Furthermore, the velocity distribution functions provide additional information that previous techniques do not. Such information may be useful, e.g., the possibility of separating mines from false targets. This new technique will be discussed, and the results from measured multibeam LDV data will be presented. [This material is based upon work supported by the U. S. Army Communications-Electronics Command Night Vision and Electronic Sensors Directorate under Contract DAAB15-02-C-0024.]

3:10–3:30 Break

3:30

1pPAa8. Rapid high spatial resolution imaging of ground vibration for buried landmine detection using ESPI. William C. Alberts II, James M. Sabatier (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, 1 Coliseum Dr., University, MS 38677, walberts@olemiss.edu), and Vyacheslav Aranchuk (MetroLaser, Inc., Irvine, CA 92614-6236)

Recent work has shown that many landmines exhibit a multimode vibration pattern. To fully map the vibration pattern of these modes requires spatial resolutions on the order of millimeters. Any practical field instrument to detect landmines that cues on such modal patterns must take advantage of parallel optical processing. An optical technique that lends itself to such vibration sensing is an electronic speckle pattern interferometer (ESPI). The double-pulsed ESPI system has been used for the vibration measurement of the ground surface. The first and second laser pulses synchronized with the vibration peak and the vibration valley, respectively, illuminate the object. The imaged object wavefronts are combined with a reference wavefront and recorded in two video frames by using a CCD camera. The spatial distribution of the 2D vibration amplitude is obtained by subtracting two speckle patterns and processing the corresponding fringe pattern. The displacement sensitivity of the ESPI is about 30 nm. Here, we not only consider airborne sound sources to excite vibrations in the ground, but also a mechanical shaker to significantly increase the vibration amplitudes at the spot of interest. The vibrational pattern at the surface of the ground over buried antitank and antipersonnel landmines is studied.

3:45

1pPAa9. Evaluation of an ultrasonic displacement sensor. Douglas Fenneman, Bradley Libbey (Night Vision and Electron. Sensors Directorate, 10221 Burbeck Rd., Fort Belvoir, VA 22060), and James Martin (Georgia Inst. of Technol., Atlanta, GA 30332)

An ultrasonic displacement sensor is currently being investigated. A possible application for this sensor is acoustic/seismic landmine detection in which a noncontact vibrometer is employed to measure the normal surface velocity of soil. The ultrasonic sensor is an alternative displacement sensor to a laser Doppler vibrometer (LDV). The ultrasonic wavelength suggests that the ultrasonic sensor is capable of resolving displacements in the presence of rough surfaces. This functionality is of particular interest because scattering of optical signals by rough surfaces can prohibit accurate LDV measurements. The ultrasonic sensor will be evaluated in an array of surface roughness conditions. The performance of the ultrasonic sensor will be compared to the LDV for each of these cases. The relative advantages and limitations of the ultrasound system will be presented. The applicability of the ultrasonic sensor to acoustic mine detection will also be addressed.

4:00

1pPAa10. Doppler ultrasonic detection of targets buried in grass-covered soil. James M. Sabatier (Univ. of Mississippi, Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38677, sabatier@olemiss.edu) and Andi G. Petculescu (Northwestern Univ., Evanston, IL 60208)

An ultrasonic Doppler vibrometer (UDV) is used outdoors to detect vibrating targets buried in grass-covered soil. The sensor head uses two solid-dielectric transducers, in a pitch–catch configuration. A first set of measurements is done using calibrated vibrational sources (shakers), whose vibrational characteristics are known and/or easily predictable. Then, the system is put to the test of detecting a landmine buried in a realistic environment. The target (landmine) is excited either by a mechanical shaker or by a loudspeaker, through acoustic-to-seismic coupling. The wind speed was monitored continuously. Since it is known that wind degrades the UDV signal, efforts were made to perform the experiments in a still environment. The UDV results are compared with those obtained with a laser Doppler vibrometer.

4:15

1pPAa11. Investigating nonlinearity in acoustic landmine experiments using a clamped-plate soil oscillator. Dang V. Duong and Murray S. Korman (Dept. of Phys., U.S. Naval Acad., Annapolis, MD 21402)

This project investigates the nonlinear effects involved in the detection of plastic landmines using the acoustic-to-seismic coupling techniques developed by Sabatier (linear methods) and Donskoy (nonlinear methods). The soil-plate oscillator is a laboratory apparatus that represents a good physical model for the VS 1.6 and VS 2.2 inert anti-tank plastic landmines. The apparatus consists of a thick-walled cylinder filled with sifted homogeneous soil resting on a thin elastic plate that is clamped to the bottom of the column. Using a loudspeaker (located over the soil) that is driven by a swept sinusoid, tuning curve experiments are performed. The vibration amplitude versus frequency is measured on a swept spectrum analyzer using an accelerometer located on the soil–air interface or under the plate. The ''backbone'' curve shows a linear decrease in peak frequency versus increasing amplitude. A two-tone test experiment is per-

formed using two loudspeakers generating acoustic frequencies (closely spaced on either side of resonance, typically ~100 Hz). A collection of combination frequency component profiles (along with the primaries) are measured across the soil surface. In particular, a double peaked profile occurred at $2f_1+f_2$ corresponding to a common timpani mode. [Work sponsored by U.S. Army Communications-Electronics Command RDEC, NVESD, Fort Belvoir, VA.]

4:30

1pPAa12. False alarms associated with the acoustic-to-seismic detection of buried land mines. Vladimir N. Fokin, Margarita S. Fokina, and James M. Sabatier (Natl. Ctr. for Physical Acoust., 1 Coliseum Dr., University, MS 38655, vfok@olemiss.edu)

Important problems in landmine detection include false alarms and clutter [high values of the acoustic-to-seismic transfer function (TF) in some frequency bands] that mimic the physics of a buried landmine. Many of these high values of the TF are due to the natural variability of the ground. In this paper both the space-frequency variability of the TF(f,x,y) and the connection with ground variability are discussed. The viscoelastic, layered model of the ground qualitatively explains high values of the TF both in certain spatial regions of the scan site and for some frequencies. This model and the measured spatial dependence of the TF(f,x,y) were used to model the space-frequency distribution of the acoustic-to-seismic transfer function. Comparison between the calculated and measured transfer functions in the space-frequency volume show that this model can satisfactorily explain variability of the acoustic-to-seismic transfer function. To investigate the spatial variability of ground parameters, the measured TF for a few sites was analyzed. This analysis reveals that different frequency modulation scales exist in the acoustic-to-seismic TF. These different frequency modulation scales are due to spatial dependencies of ground parameters and ground layering. [Work supported by ONR Grant N00014-02-1-0878.]

MONDAY AFTERNOON, 24 MAY 2004 CONFERENCE ROOM E, 1:00 TO 5:30 P.M.

Session 1pPAb

Physical Acoustics and Biomedical Ultrasound/Bioresponse to Vibration: Sono Et Gravitas: Robert E. Apfel Memorial Session

Lawrence A. Crum, Cochair
Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105-6698

Ronald A. Roy, Cochair
Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Chair's Introduction—1:00

Invited Papers

1:10

1pPAb1. Cavitation nucleation. Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

For his dissertation research at Harvard, Bob Apfel chose the subject of homogeneous nucleation, and conceived of some ingenious experiments to test existing theories. By selecting a small microdroplet of liquid, he could make the reasonable assumption that no inhomogeneities were present to serve as preferential sites for liquid rupture. However, Bob also studied dirty liquids, as well as very clean ones, and wrote some seminal papers on inhomogeneous nucleation, in which he developed the Golden rule: Know thy liquid! Currently, considerable attention has been devoted to the study of cavitation generation *in vivo*, particularly in blood, and, for this case, the nucleation conditions are much different than those for normal liquids. In this presentation, I will review some of Bob's pioneering studies and present some of our latest studies of cavitation inception, both *in vitro* and *in vivo*.

1:25

1pPAb2. The oscillation of vapor bubbles. Andrea Prosperetti and Zhizhong Yin (Dept. of Mech. Eng., Johns Hopkins Univ., Baltimore, MD 21218)

Bob Apfel had so many interests that it is impossible—however fitting and desirable—to pay homage to his work as a whole. Some of his early studies were devoted to bubble nucleation at high superheats. In the first part of this paper a recent application of this phenomenon is described. Once a vapor bubble is generated, its subsequent oscillations (free and forced) present analogies and differences with those of a gas bubble: the second part of the paper focuses on this topic. [Work supported by NSF and NASA.]

1:40

1pPAb3. Inspection of the interior of a collapsing bubble via molecular dynamics. Werner Lauterborn, Thomas Kurz, Burkhard Metten, and Daniel Schanz (Drittes Physikalisches Institut, Universitaet Goettingen, Buergerstr. 42-44, D-37073 Goettingen, Germany, lb@physik3.gwdg.de)

Strongly collapsing bubbles emit shock waves, faint light flashes and induce chemical reactions. At present, the interior of a collapsing bubble is not yet accessible experimentally to reveal the processes behind this behavior. Thus our knowledge has to be advanced theoretically. The standard method relies on solving a set of partial differential equations. However, in the case of small sonoluminescent bubbles, the number of molecules may become too small for a continuum approach. Therefore, the processes within collapsing sonoluminescent bubbles are investigated by molecular dynamics simulations. The temperature, density and pressure distribution within the bubble are calculated. Whereas mass and heat diffusion inside the bubble are automatically accounted for, they must be explicitly introduced across the boundary to the liquid. In particular, also the dissociation of water vapor and chemical reactions of the dissociation products are taken into account. Results are presented for different acoustic driving conditions (sound pressure amplitude and frequency) and different gas compositions including mixtures of noble gases and water vapor. A prediction of the light emission is given to connect the calculations with observable parameters.

1:55

1pPAb4. Quantitative theoretical explanation of Apfel's experimental phase diagrams for sonoluminescing bubbles. Detlef Lohse and Ruediger Toegel (Dept. of Phys., Univ. of Twente, 7500 AE Enschede, The Netherlands)

Robert Apfel had an enormous impact on the research on single bubble sonoluminescence, the light emission of a single sound driven bubble [for a review, see Brenner *et al.*, Rev. Mod. Phys. **74**, 425 (2002)]. In 1996, at the ASA Meeting in Hawaii, he posed a challenge to the theoreticians in the field: Make experimentally testable predictions on single bubble sonoluminescence. Apfel collected the predictions and gave a wonderful review talk on the state of the field. Later, he several times came back to that list, comparing the predictions with latest experimental results. Our own predictions those days referred to the phase diagrams of single bubble sonoluminescence. Later Apfel himself, together with Ketterling, measured those phase diagrams experimentally [J. A. Ketterling and R. E. Apfel, Phys. Rev. Lett. **81**, 4991 (1998)]. Though qualitatively our 1996 predictions turned out to be correct, a full quantitative model could only be developed recently [R. Toegel and D. Lohse, J. Chem. Phys. **118**, 1863–1875 (2003)]. In the presentation we will compare the model predictions with Apfel's data.

2:10

1pPAb5. From blood to bubbles: Time resolved micro-particle detection and characterization by scattered ultrasound. Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ronroy@bu.edu)

Robert Apfel believed in the creative application of acoustics technology to difficult problems in biomedical sensing. Much of his work in this area focused on material characterization, with the intention of effecting diagnosis. His early work in blood cell characterization employed acoustic levitation to measure the bulk mechanical properties of human red blood cells. This subsequently paved the way to the use of high-frequency acoustic scattering to yield the compressibility and density of individual blood cells. Technology developed in this later effort was then adapted to the very difficult problem of transient micro-cavitation detection, and the active cavitation detector (ACD) was born. This paper traces this line of work from its origins and, in the process, serves to celebrate Bob Apfel's peerless ingenuity and irrepressible creativity.

2:25

1pPAb6. Robert Apfel's contribution to clinical diagnostic ultrasound: The mechanical index. Christy K. Holland (Dept. of Biomed. Eng., 234 Albert Sabin Way, Cincinnati, OH 45267-0586)

The mechanical index, MI, resulted from theoretical considerations of the short-pulse acoustic threshold for inertial cavitation in water populated with microbubbles of all sizes [R. E. Apfel and C. K. Holland, Ultrasound Med Biol. **17**, 179–185 (1991)]. In this review, the onset of cavitation will be discussed with reference to Robert Apfel's legacy of theoretical and experimental data. The questions arise: Can the utility of the MI be extended to situations in which the threshold MI is exceeded, thereby allowing for some estimate of the quantification of a potential bioeffect due to microcavitation? Also, can the MI be extended to situations in which pulses are, unlike the original formulation, not short? Is there a theoretical or semi-empirical basis for the MI threshold below which cavitation is unlikely? Can the MI be used to predict gas contrast agent destruction? The possible consequences of gas body activation associated with aerated lung tissue, intestinal gas pockets or encapsulated gas contrast agents represent specific instances of cavitation considerations relevant to clinical practice. Monitoring the real-time display of the MI (mandated by the FDA) helps clinicians evaluate and minimize the potential risks in the use of diagnostic ultrasound instrumentation. [Research supported by National Institutes of Health Grant R29 HL58761.]

2:40

1pPAb7. Active cavitation detection of asymmetrical inertial cavitation. E. Carr Everbach (Dept. of Eng., Swarthmore College, Swarthmore, PA 19081)

The active cavitation detector (ACD) developed in Bob Apfel's laboratory has often been employed to quantify pressure thresholds for inception of symmetrical inertial cavitation of microbubbles. In the current application, however, a 30-MHz ACD interrogates individual echo-contrast agent bubbles adhering to a Mylar(TM) sheet that are driven into asymmetrical (jet-producing) collapse by a 1-MHz toneburst (>1 MPa pp). The resulting ACD output suggests that asymmetrical bubble collapse is slower than symmetrical

collapse, producing less total radiated acoustic power. ACD output mixed with reference sinusoids at 30 MHz and low pass filtered yields Doppler signals that may be useful in quantifying asymmetrical collapses under biomedically relevant conditions, such as on endothelial walls.

2:55

1pPAb8. Microlesions induced by microcavitation during contrast echocardiography. Douglas Miller, Peng Li, David Gordon, and William Armstrong (Univ. of Michigan Medical Ctr., Ann Arbor, MI 48109, douglm@umich.edu)

The purpose of this study was to search for histologically identifiable lesions associated with myocardial contrast echocardiography (MCE) in rats. Diagnostic ultrasound scans with 1:4 end-systolic triggering provided a short-axis view of the left ventricle in rats at 1.5 MHz with 1.45-μs pulses of 1.7 Mechanical Index. Two relatively high doses (500 μl/kg) of Optison^{TM} ultrasound contrast agent were given 5 min apart during 10 min of MCE. One day after scanning, rats were sacrificed and the hearts fixed for histology. Slides were scored blind by a pathologist, and photomicrographs in the anterior half of the heart sections were characterized by digital image analysis. Microlesions identified by inflammatory infiltrates were scattered primarily over the anterior half of the sections. Pathologically, there was inflammatory cell infiltration in areas of $0.6 \pm 0.5\%$ of the sections for shams and $3.6 \pm 3.6\%$ for MCE ($P < 0.01$). Analysis of the photographs from the anterior wall found microlesion areas of $0.5 \pm 0.8\%$ for shams and $7.4 \pm 5.0\%$ for MCE ($P < 0.02$). Diagnostic MCE at high Mechanical Index has a potential for causing microscale lesions in the myocardium by nucleation of microcavitation. [Work supported by NIH Grant EB0338.]

3:10–3:25 Break

3:25

1pPAb9. Shape oscillations of acoustically levitated drops in water: Early research with Bob Apfel on modulated radiation pressure. Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814)

In 1976, research in collaboration with Bob Apfel demonstrated that low-frequency shape oscillations of hydrocarbon drops levitated in water could be driven using modulated radiation pressure. While that response to modulated ultrasound was subsequently extended to a range of systems, the emphasis here is to recall the initial stages of development in Bob Apfel's laboratory leading to some publications [P. L. Marston and R. E. Apfel, J. Colloid Interface Sci. **68**, 280–286 (1979); J. Acoust. Soc. Am. **67**, 27–37 (1980)]. The levitation technology used at that time was such that it was helpful to develop a sensitive method for detecting weak oscillations using the interference pattern in laser light scattered by levitated drops. The initial experiments to verify this scattering method used shape oscillations induced by modulated electric fields within the acoustic levitator. Light scattering was subsequently used to detect shape oscillations induced by amplitude modulating a carrier having a high frequency (around 680 kHz) at a resonance of the transducer. Methods were also developed for quantitative measurements of the drop's response and with improved acoustic coupling drop fission was observed. The connection with research currently supported by NASA will also be noted.

3:40

1pPAb10. Surface-controlled drop oscillations in space. R. Glynn Holt (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215)

A series of experiments probing the effects of surfactants was performed by Bob Apfel and his research group in the 1990s. Several laboratory experiments were carried out in uni-axial acoustic levitators. Two experiments were carried out in a triple-axis levitator called the Drop Physics Module, which was carried on Space Shuttle Columbia as part of the First and Second United States Microgravity Laboratory missions. Liquid drops containing aqueous solutions of soluble surfactants were acoustically positioned and deformed (and in some cases rotated) in order to excite shape mode oscillations. The results of these experiments allowed the inference of surface rheological properties (Gibb's elasticity, surface viscosity coefficients) as functions of surfactant type and concentration. The highlights of this effort will be presented in a semi-technical fashion. [Work supported by NASA.]

3:55

1pPAb11. Apfel's superheated drop detector. Francesco d'Errico (Dept. of Therapeutic Radiol., Yale Univ. School of Medicine, HRT-219, P.O. Box 208040, New Haven, CT 06520-8040)

The introduction of new approaches for radiation dosimetry is rare. A similar breakthrough occurred in 1979, when Robert Apfel invented the superheated drop detector, a miniature relative of the bubble chamber. A fundamental in high-energy particle physics, the bubble chamber utilizes a liquid briefly brought to a transient, radiation-sensitive superheated state by reducing its pressure. Mass boiling of the liquid is prevented by cyclic pressurization, drastically limiting the detection efficiency. In Apfel's detector, the liquid is kept in a steady superheated state by fractionating it into droplets and dispersing them in an immiscible host fluid, a perfectly smooth and clean container. The approach extends the lifetime of the metastable droplets to the point that practical application in radiation dosimetry is possible. Bubble formation is measured from the volume of vapor or by detecting individual vaporizations acoustically. Various halocarbons are employed and this permits a wide range of applications. Moderately superheated halocarbons are used for neutron measurements, since they are only nucleated by energetic neutron recoil particles. Highly superheated halocarbons nucleate with much smaller energy deposition and are used to detect photons and electrons. This paper reviews the radiation physics of superheated emulsions and their manifold applications.

4:10

1pPAb12. Testing thin film adhesion strength acoustically. Sameer I. Madanshetty (Kansas State Univ., Manhattan, KS 66506, sameer@ksu.edu), Kevin M. Wanklyn, and Hang Ji (Uncopiers Inc., Manhattan, KS 66506)

A new method of measuring the adhesion strength of thin films to their substrates is reported. The method is based on an analogy with the common tensile test of materials. This is an acoustic method that uses acoustic microcavitation to bring about controlled erosion of the thin film. Based on the insonification pressure and the time to complete erosion, the adhesion strength is assessed. The measurements correctly rank order a set of thin film samples of known adhesion strengths.

4:25

1pPAb13. Fabrication of PLGA polymer microspheres for U. S. mediated gene delivery. Rene G. Williamson, William M. Saltzman (Dept. of Biomed. Eng., Yale Univ., 15 Prospect St., Becton 225, New Haven, CT 06511, rgw27@pantheon.yale.edu), and Janet L. Brandsma (Yale Univ., New Haven, CT 06520)

The promises of gene therapy remain unfulfilled because of the lack of a safe and efficient method for transfecting DNA into cells. PLGA has been used as a vehicle for protein, drug, and gene delivery applications because of its biocompatibility and sustained release properties. PLGA polymer microspheres offer advantages of safety and the possibility of sustained intracytoplasmic delivery. The PLGA also protects the plasmid from degradation. Using the double-emulsion microsphere fabrication technique, a new DNA delivery vehicle, comprising of plasmid DNA and octafluoropropane gas encapsulated in PLGA polymer and PVA stabilizer (Sonospheres) was made. The encapsulated gas offers acoustic activity to the microspheres, which enables them to undergo cavitation in an acoustic field. The goal is to lead to increased DNA transfection when these Sonospheres are subjected to an acoustic field in the MHz frequency range. A summary of the fabrication methods and some initial *in vitro* studies will be presented.

4:40

1pPAb14. Reflector geometry and pressure at the focus of a shock wave lithotripter. Jonathan Iloreta, Prahallad Iyengar, and Andrew Szeri (Dept. of Mech. Eng., Univ. of California, Berkeley, Berkeley, CA 94720-1740, aszeri@me.berkeley.edu)

It has been experimentally shown [Zhong, J. Acoust. Soc. Am. **113**, 586–597 (2003)] that refinement of the geometry of the HM-3 lithotripter reflector suppressed the expansion of cavitation bubbles without compromising stone comminution. This effect has been attributed to a change in the rarefaction tail of the pressure wave near the second focus of the original reflector (F_2). Following this idea, a numerical model of the reflection and steepening of a pressure wave from an axisymmetric lithotripter has been developed. The model is based on the Euler equations coupled with the Tait equation of state. Preliminary results of the pressure fields produced by numerous reflector shapes are presented. The results show the changes in the rarefaction tail of the pressure wave near F_2 for different geometries, thus hinting at the possibility of optimization of the reflector shape for an ideal waveform. [Work supported by NSF.]

4:55

1pPAb15. Reduction of variance in statistical estimates of spectra for use in correction of ultrasonic aberration. Jeffrey P. Astheimer, Wayne C. Pilkington, and Robert C. Waag (Ultrasound Res. Lab., Univ. of Rochester Med. Ctr., Box 648, Rochester, NY 14642)

Parameters of ultrasonic aberration can be obtained from power spectra of scattering when individual scattering measurements from which the spectra are estimated have a common aberration and the same nominal geometry. However, the scattering volumes are then confined to a small spatial region and use of finitely many overlapping volumes that result in a nonzero variance is necessary for the measurements. Assuming the scattering is from a spatially uncorrelated medium, the variance of the spectral estimates is expressed as the product of the variance for a single measurement and a reduction factor that depends on the amount of overlap between each volume pair. This factor describes the rate of convergence and the accuracy of the estimates as a function of the number and the overlap of the scattering volumes. Assuming further that the individual volumes are localized by a Gaussian window and that the centers of the volumes are located on orbits of an icosahedral rotation group, the variance is minimized by adjusting the weight and radius of each orbit. Numerical evaluations using orbits formed by icosahedral vertices, face centers, and edge midpoints show that a significant reduction of variance can be achieved from volumes in a confined region.

5:10–5:30
Announcement
Acoustics Research Letters Online (ARLO)
Special Apfel Memorial Edition

Session 1pPP

Psychological and Physiological Acoustics: Compression in Hearing

Sid P. Bacon, Chair
Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287-1908

Chair's Introduction—2:15

Invited Papers

2:20

1pPP1. Compression in the cochlea. Mario A. Ruggero (Hugh Knowles Ctr., Dept. of Commun. Sci., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208-3550, mruggero@northwestern.edu)

Cochlear compression consists of the reduction of the amplitude range of the auditory signal, from 6 orders of magnitude at the stapes to (typically) 1–2 orders of magnitude in the spike rate of individual cochlear afferents. Psychophysical discussions of compression generally focus on basilar-membrane (BM) vibrations at the base of the cochlea, which at the characteristic frequency (only) grow with stimulus intensity at remarkably compressive rates but exhibit almost negligible harmonic distortion. In apical cochlear regions, however, BM compression appears to be substantially weaker than at the base, involves significant dc and harmonic distortion, and is not confined to frequencies near the characteristic frequency. Additional stages of compression exist in mechanical-to-electrical transduction in inner hair cells and in the spike generation mechanism. Both the regional variations in the features of BM nonlinearities and the additional compression introduced by proximal stages of signal transformation should be addressed in physiology-based psychophysical models of auditory processing. Ideally, such models should take into account off-frequency listening and should be based on the output of the cochlea, i.e., the activity of the entire array of cochlear afferents, which is better known and more relevant to central auditory processing than BM vibrations. [Work supported by NIH.]

2:45

1pPP2. Behavioral estimates of compression in normal and impaired ears. Christopher J. Plack (Dept. of Psych., Univ. of Essex, Wivenhoe Park, Colchester CO4 3SQ, England, cplack@essex.ac.uk) and Andrew J. Oxenham (MIT, Cambridge, MA 02139-4307)

Over the last 7 years, forward masking techniques have been used to estimate peripheral compression in the human auditory system. The growth of masking and temporal masking curve (TMC) techniques estimate the response to a tone at characteristic frequency (CF) by comparing the masking function (masker level as a function of signal level and masker-signal interval, respectively) for a masker at the signal frequency with the masking function for a masker below the signal frequency (assumed to be processed linearly at high CFs). Compression can also be estimated using the additivity of forward masking technique, in which the effects on signal threshold of combining two equally effective maskers are used to derive the exponent. Overall, the results suggest strong compression in humans (exponent of 0.2) across the range of CFs tested so far (250 to 8000 Hz). The behavioral techniques also suggest that a sensorineural hearing loss of greater than about 50 dB results in an almost complete linearization of the response. However, recent TMC data suggest that less severe hearing losses are associated with a reduction in gain, and a reduction in the level range over which compression is present, but no reduction in the maximum compression.

3:10

1pPP3. Perceptual consequences of normal and abnormal peripheral compression: Potential links between psychoacoustics and speech perception. Andrew J. Oxenham, Peninah S. Rosengard, and Louis D. Braida (Res. Lab. of Electron., MIT, Cambridge, MA 02139, oxenham@mit.edu)

Cochlear damage can lead to a reduction in the overall amount of peripheral auditory compression, presumably due to outer hair cell (OHC) loss or dysfunction. The perceptual consequences of functional OHC loss include loudness recruitment and reduced dynamic range, poorer frequency selectivity, and poorer effective temporal resolution. These in turn may lead to a reduced ability to make use of spectral and temporal fluctuations in background noise when listening to a target sound, such as speech. We tested the effect of OHC function on speech reception in hearing-impaired listeners by comparing psychoacoustic measures of cochlear compression and sentence recognition in a variety of noise backgrounds. In line with earlier studies, we found weak (nonsignificant) correlations between the psychoacoustic tasks and speech reception thresholds in quiet or in steady-state noise. However, when spectral and temporal fluctuations were introduced in the masker, speech reception improved to an extent that was well predicted by the psychoacoustic measures. Thus, our initial results suggest a strong relationship between measures of cochlear compression and the ability of listeners to take advantage of spectral and temporal masker fluctuations in recognizing speech. [Work supported by NIH Grants Nos. R01DC03909, T32DC00038, and R01DC00117.]

3:35

1pPP4. Effect of compressor design on auditory function and a psychoacoustic Turing test. Brent W. Edwards (Sound ID, 3430 W. Bayshore Rd., Palo Alto, CA, brent@edwards.net)

Hearing aids incorporate multiband compression to compensate for the loudness recruitment that results from sensorineural hearing impairment. No consensus exists in the hearing-aid industry on the best compressor design or on what the design criteria should be. Differences exist in compressor time constants, number of bands, filter shapes, and fitting formula. Such design differences can result in different aided auditory ability as measured by psychoacoustic tests. This research investigated the effect of different compressor designs on fundamental psychoacoustic ability. Multiband compressors with different time constants and different analysis/synthesis filterbanks designs were simulated. Aided performance by the hearing impaired was calculated under each configuration for several psychoacoustic measures, including forward masking, loudness summation, and simultaneous off-frequency masking. Differences in aided performance in these tasks were found: some compressor designs produced aided psychoacoustic ability that resembled normal performance, while other compressor designs did not significantly alter performance from unaided impairment. A psychoacoustic Turing test is proposed that incorporates these aided performance measures as part of a hearing-aid design, a validation strategy, and a comparative technique for evaluating different hearing-aid designs.

4:00–4:15 Break

4:15

1pPP5. Compression, cochlear implants, and psychophysical laws. Fan-Gang Zeng (364 Med. Surge. II, Univ. of California, Irvine, CA 92612)

Cochlear compression contributes significantly to sharp frequency tuning and wide dynamic range in audition. The physiological mechanism underlying the compression has been traced to the outer hair cell function. Electric stimulation of the auditory nerve in cochlear implants bypasses this compression function, serving as a research tool to delineate the peripheral and central contributions to auditory functions. In this talk, I will compare psychophysical performance between acoustic and electric hearing in intensity, frequency, and time processing, and pay particular attention to the data that demonstrate the role of cochlear compression. Examples include both the cochlear-implant listeners' extremely narrow dynamic range and poor pitch discrimination and their exquisite sensitivity to changes in amplitude and phase. A unified view on the complementary contributions of cochlear compression and central expansion will be developed to account for Webers' law and Stevens power law.

Contributed Papers

4:40

1pPP6. The compression curves of different models of cochlear nonlinearities. James M. Harte, Stephen J. Elliott (Inst. of Sound and Vib. Res., Univ. of Southampton, Southampton SO17 1BJ, UK), and Henry Rice (Trinity College, Dublin, Ireland)

A widely used method for representing the nonlinear response of the basilar membrane (BM) is a graph of log amplitude of the BM motion against the log amplitude of the sinusoidal driving pressure, known as the input–output level curve. At low sound pressure levels (less than approximately 30 dB) the level of the BM response rises with sound pressure level at a slope of about 1 dB/dB, indicating a linear response. Above this region the slope of the level curve decreases, typically to about 1/2 to 1/3 dB/dB, indicating a compressive nonlinearity. Various models for cochlear nonlinearity will be presented, particularly contrasting the characteristics and behavior of an instantaneously acting nonlinear function and those of level-dependent systems. Both of these models have been used historically to model the nonlinearities present in the cochlea. Comparisons will be made between the properties of these models, highlighting potential methods of distinguishing between them. Similar nonlinear responses to those of the BM are observed in otoacoustic emissions. Experimental results will be reported which attempt to distinguish between the instantaneously acting and level-dependent models of this nonlinearity, which suggest that the latter is more likely.

4:55

1pPP7. The temporal effect in a notched-noise masker for normal-hearing and hearing-impaired listeners. Elizabeth A. Strickland (Dept. of Audiol. and Speech Sci., Purdue Univ., West Lafayette, IN 47907, estrick@purdue.edu)

This is another in a series of studies on the relationship between the temporal effect (or overshoot) and the cochlear active process. The temporal effect refers to the change in threshold for a short-duration tone as it is delayed from the onset of a longer-duration masker. In a previous study, the temporal effect was measured as a function of signal level in a broadband masker, for listeners with normal hearing and listeners with varying degrees of cochlear hearing loss. When analyzed in terms of cochlear input–output functions, the results were consistent with a decrease in gain at the signal frequency during the course of the masker, as well as a decrease in gain with hearing impairment. In the present study, the temporal effect was measured for a 10-ms, high-frequency signal in a masker with a symmetric notch (relative frequency of 0.3) about the signal frequency, for listeners with normal hearing and cochlear hearing loss. This condition is of interest because the thresholds may reflect the effects of suppression as well as excitatory masking. The relationship between the degree of hearing impairment and the growth of masker level with signal level will be analyzed. [Work supported by the Kinley Trust.]

5:10

1pPP8. Cochlear nonlinearity between 500 and 8000 Hz in listeners with moderate cochlear hearing loss. Enrique A. Lopez-Poveda (Instituto de Neurociencias de Castilla y Leon, Universidad de Salamanca, Avda. Alfonso X El Sabio s/n, Salamanca, Spain, ealopezpoveda@usal.es), Christopher J. Plack, Ray Meddis (Univ. of Essex, Colchester C04 3SQ, UK), and Jose L. Blanco (Oticon Espaa, 28108 Alcobendas, Madrid, Spain)

For two listeners with flat hearing loss, temporal masking curves (TMCs) were measured for probe frequencies (Fp) of 0.5, 1.0, 2.0, 4.0, and 8.0 kHz and masker frequencies of 0.5, 0.6, 0.7, 0.9, 1.0, 1.05, 1.1, and 1.2 Fp. From these, basilar membrane (BM) input/output (IO) functions and psychophysical tuning curves (PTCs) were derived and compared with corresponding data for normal-hearing listeners. Linear IO functions were observed for one ear only and for probe frequencies of 4 kHz and above. This result is consistent with the consequences of severe outer hair cell (OHC) damage on the BM response. Elsewhere, however, IO functions showed residual compression, with slopes in the compression region close to normal. This suggests that either the impairment relates mostly to inner hair cell (IHC) dysfunction, or that OHC dysfunction reduces the gain of the BM at low levels while maintaining the degree of maximum compression. The data suggest that the active and the passive

BM mechanisms have different relative positions (in frequency) in the apical and the basal regions of the cochlea. Across probe frequencies, the slopes of the TMCs for off-frequency maskers are shallower for the impaired ears. It is considered that this possibly reflects IHC-related compression.

5:25

1pPP9. A hearing-aid signal-processing scheme based on the temporal aspects of compression. Laurel H. Carney (Inst. for Sensory Res. and Dept. of Bioeng. and Neurosci., Syracuse Univ., Syracuse, NY, lacarney@syr.edu), Lufeng Shi, and Karen A. Doherty (Syracuse Univ., Syracuse, NY)

Changes in gain associated with the basilar membrane compressive nonlinearity are accompanied by changes in the bandwidth of tuning. Filters with level-dependent bandwidth have level-dependent phase properties. These phase properties result in level-dependent timing of sustained phase-locked responses of auditory-nerve (AN) fibers at low frequencies and level-dependent latencies at high frequencies, where phase-locking rolls off. In the healthy ear, level-dependent temporal aspects of AN responses carry information about stimulus level and spectral properties. Loss of compression with hearing impairment thus results not only in a reduction of amplification, but also in distortion of the temporal response pattern of the AN. The temporal aspects of compression suggest that signal-processing schemes that attempt to correct sounds, or restore normal spatio-temporal response patterns, should include dynamic level-dependent phase shifts. A nonlinear signal-processing scheme will be presented which includes dynamic frequency- and level-dependent phase shifts, based on physiological models of the temporal response properties of AN fibers. Preliminary testing measured listeners preferences for sentences and intelligibility of vowel-consonant syllables with different degrees of nonlinear processing. Hearing-impaired listeners tended to prefer the dynamically corrected stimuli based on improved clarity. Correction also improved intelligibility for some phonemes. [Work supported by NIDCD R21-006057.]

5:40

1pPP10. The effect of amplitude compression, time constants, and number of channels on speech intelligibility in noise. Rolph Houben and Guido F. Smoorenburg (UMC-Utrecht, Heidelberglaan 100, 3508 GA Utrecht, The Netherlands)

The influence of several compression parameters on speech intelligibility in speech-shaped noise was systematically investigated. Experimental conditions include all combinations of compression ratio (CR=1/2, 2/2, 2/3, 3/3 for low/high frequencies), attack and release times (Ta/Tr =4/4, 4/40, 40/40, 4/400, and 40/400 ms), and number of channels (1, 2, and 6). Twenty subjects with moderate sensorineural hearing loss took part in the experiment. The best average speech reception threshold occurred with two-channel compression at a compression ratio of 2/3 and Ta/Tr =40/40 ms. It was 0.7 dB better than linear amplification. The standard deviation between subjects was 1.7 dB. With six-channel compression the best results (0.4 dB better than linear amplification) were obtained at larger time constants (Ta/Tr=40/400 ms) and CR=2/2 and 2/3. The best result with single-channel compression was equal to linear amplification and occurred for the longest time constant (Ta/Tr=40/400 ms) and CR =2/2. In fluctuating noise the best speech reception threshold was found with single-channel compression at CR=2/2 and Ta/Tr=4/400 ms. It was 0.9 dB better than linear amplification; the standard deviation was 2.3 dB. The results suggest that finding the optimal condition for an individual is tedious.

5:55–6:30

Panel Discussion

MONDAY AFTERNOON, 24 MAY 2004 IMPERIAL BALLROOM B, 1:00 TO 5:00 P.M.

Session 1pSC

Speech Communication: Poster Session II

Carole E. Gelfer, Chair

Department of Communication Disorders, William Patterson University, Wayne, New Jersey 07470

Contributed Papers

All posters will be on display from 1:00 p.m. to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 5:00 p.m.

1pSC1. Identification of 20-ms speech samples from normal and neurologically impaired talkers: Do coarticulation differences matter? Joan E. Sussman, Kris Tjaden, and Ya-ju Yu (Dept. of Commun. Disord. and Sci., Univ. at Buffalo, 122 Cary Hall, 3435 Main St., Buffalo, NY 14214)

Prior studies of speech suggest that the extent of coarticulation varies both across and within talkers. Speakers sometime use average, high, or low amounts of coarticulation. The perceptual consequences of coarticulatory variability are not well understood, particularly for the speech of individuals with motor control disorders. In the current study, listeners heard the same 20-ms tokens that were used to calculate the degree of anticipatory coarticulation of the following vowel. Acoustic measures were made from the first 20 ms of stop consonants [k] or [t], or from a point 70 ms preceding the vowel in [s] productions. Healthy talkers, speakers with Multiple Sclerosis, and speakers with Parkinson's disease produced the stimuli. Speech tokens included three degrees of anticipatory coarticulation: average, high, and low. The listeners' task was to identify whether the vowels that followed the consonant were originally [i] or [u]. Results showed that identification accuracy was poorer for the current 20-ms tokens than for prior speech samples including all aperiodicity of

consonants. Identification was poorest for [s] and best for [k] tokens. The speed of listener response gave some indication of the differing extents of coarticulation. Perceptual validity for the 20-ms measure of coarticulation is discussed.

1pSC2. Are vowel errors influenced by consonantal context in the speech of persons with aphasia? Carole E. Gelfer (Dept. of Commun. Disord., William Paterson Univ., Wayne, NJ 07470, gelferc@wpunj.edu), Fredericka Bell-Berti (St. John's Univ., Jamaica, NY 11439), and Mary Boyle (Montclair State Univ., Upper Montclair, NJ 07043)

The literature suggests that vowels and consonants may be affected differently in the speech of persons with conduction aphasia (CA) or nonfluent aphasia with apraxia of speech (AOS). Persons with CA have shown similar error rates across vowels and consonants, while those with AOS have shown more errors for consonants than vowels. These data have been interpreted to suggest that consonants have greater gestural complexity than vowels. However, recent research [M. Boyle *et al.*, Proc. International Cong. Phon. Sci., 3265–3268 (2003)] does not support this interpretation: persons with AOS and CA both had a high proportion of vowel errors, and vowel errors almost always occurred in the context of consonantal errors. To examine the notion that vowels are inherently less complex than consonants and are differentially affected in different types of aphasia, vowel production in different consonantal contexts for speakers with AOS or CA was examined. The target utterances, produced in carrier phrases, were bVC and bV syllables, allowing us to examine whether vowel production is influenced by consonantal context. Listener judgments were obtained for each token, and error productions were grouped according to the intended utterance and error type. Acoustical measurements were made from spectrographic displays.

1pSC3. The effects of complementary and alternative medicine on the speech of patients with depression. Michael Fraas (Dept. of Commun. Sci. and Disord., Univ. of New Hampshire, Durham, NH 03824, michael.fraas@unh.edu) and Michele Solloway (Univ. of New Hampshire, Durham, NH 03824)

It is well documented that patients suffering from depression exhibit articulatory timing deficits and speech that is monotonous and lacking pitch variation. Traditional remediation of depression has left many patients with adverse side effects and ineffective outcomes. Recent studies indicate that many Americans are seeking complementary and alternative forms of medicine to supplement traditional therapy approaches. The current investigation wishes to determine the efficacy of complementary and alternative medicine (CAM) on the remediation of speech deficits associated with depression. Subjects with depression and normal controls will participate in an 8-week treatment session using polarity therapy, a form of CAM. Subjects will be recorded producing a series of spontaneous and narrative speech samples. Acoustic analysis of mean fundamental frequency ($F0$), variation in $F0$ (standard deviation of $F0$), average rate of $F0$ change, and pause and utterance durations will be conducted. Differences pre- and post-CAM therapy between subjects with depression and normal controls will be discussed.

1pSC4. The effects of Parkinson's disease on the production of contrastive stress. Henry S. Cheang and Marc D. Pell (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine W., Montreal, QC H3G 1A8, Canada, hchean@po-box.mcgill.ca)

Reduced speech intelligibility has been observed clinically among patients with Parkinson's disease (PD); one possible contributor to these problems is that motor limitations in PD reduce the ability to mark linguistic contrasts in speech using prosodic cues. This study compared acoustic aspects of the production of contrastive stress (CS) in sentences that were elicited from ten subjects with PD and ten matched control subjects without neurological impairment. Subjects responded to questions that biased them to put emphasis on the first, middle, or last word of target utterances. The mean vowel duration and mean fundamental frequency ($F0$) of each keyword were then measured, normalized, and analyzed for possible differences in the acoustic cues provided by each group to signal emphatic stress. Both groups demonstrated systematic differences in vowel lengthening between emphasized and unemphasized words across word positions; however, controls were more reliable than PD subjects at modulating the $F0$ of emphasized words to signal its location in the utterance. Group differences in the $F0$ measures suggest one possible source of the impoverished intelligibility of Parkinsonian speech and will be investigated in a subsequent study that looks at the direct impact of these changes on emphasis perception by listeners. [Work supported by CIHR.]

1pSC5. Phonological processing among good and poor readers. Ratree Wayland (Prog. in Linguist., Univ. of Florida, 4131 Turlington Hall, Gainesville, FL 32611-5454, ratree@ufl.edu)

Many researchers believe that a connection exists between phonological processing skills and reading ability, and phonological deficits have often been cited as possible explanation for reading disability among both children and adults. This study will present research findings on phonological processing of various speech sounds among school-aged children who were classified as good and poor readers by standardized tests. These subjects will be administered speech discrimination tests using a variety of speech stimuli. Results of their performance on these tasks will be presented and a relationship between their reading and phonological processing abilities will be discussed.

1pSC6. Discrimination and identification of long vowels in children with typical language development and specific language impairment. Hia Datta, Valerie Shafer (Speech and Hearing Sci., The Grad. Ctr., CUNY, 365 Fifth Ave., New York, NY 10016), and Diane Kurtzberg (Albert Einstein College of Medicine, Yeshiva Univ., Bronx, NY)

Researchers have claimed that children with specific language impairment (SLI) have particular difficulties in discriminating and identifying phonetically similar and brief speech sounds (Stark and Heinz, 1966; Studdert-Kennedy and Bradley, 1997; Sussman, 1993). In a recent study (Shafer *et al.*, 2004), children with SLI were reported to have difficulty in processing brief (50 ms), phonetically similar vowels (/I-E/). The current study investigated perception of long (250 ms), phonetically similar vowels (/I-E/) in 8- to 10-year-old children with SLI and typical language development (TLD). The purpose was to examine whether phonetic similarity in vowels leads to poorer speech-perception in the SLI group. Behavioral and electrophysiological methods were employed to examine discrimination and identification of a nine-step vowel continuum from /I/ to /E/. Similar performances in discrimination were found for both groups, indicating that lengthening vowel duration indeed improves discrimination of phonetically similar vowels. However, these children with SLI showed poor behavioral identification, demonstrating that phonetic similarity of speech sounds, irrespective of their duration, contribute to the speech perception difficulty observed in SLI population. These findings suggest that the deficit in these children with SLI is at the level of working memory or long term memory representation of speech.

1pSC7. Within-speaker speech intelligibility in dysarthria: Variation across a reading passage. Yana Yunusova, Gary Weismer, John Westbury, and Nicole Rusche (Dept. of Commun. Disord. and Waisman Ctr., Univ. of Wisconsin—Madison, 1500 Highland Ave., Madison, WI 53705)

Understanding factors underlying intelligibility deficits in dysarthria is important for clinical and theoretical reasons. Correlation/regression analyses between intelligibility measures and various speech production measures (e.g., acoustic or phonetic) are often reported in the literature. However, the analyses rarely control for the effect of a third variable

(severity of speech disorder, in this case) likely to be correlated with the primary correlated variables. The current report controls for this effect by using a within-speaker analysis approach. Factors that were hypothesized to underlie the intelligibility variations in multiple breath groups within a connected discourse included structural elements (e.g., number of total words) as well as acoustic measures (e.g., $F2$ variation). Results showed that speech intelligibility in dysarthric speakers with two forms of neurological disease (Parkinson and ALS) does, in fact, vary across breath groups extracted from a connected discourse, and that these variations are related in some cases to a per breath estimate of $F2$ variation. [Work supported by NIDCD Award No. R01 DC03723.]

1pSC8. Speech prosody in Friedreich's and olivo-ponto cerebellar atrophy. Maureen Casper (Grad. School and Univ. Ctr. of City Univ. of New York, NY)

A critical issue in the study of speech motor control is the identification of the mechanisms that generate the temporal flow of serially ordered articulatory events. Two staged models of serial ordered events (Lashley, 1951; Lindblom, 1963) claim that time controls events whereas dynamic models predict a relative relation between time and space. Each of these models predicts a different relation between the acoustic measures of formant frequency and segmental duration. The most recent method described herein provides a sensitive index of speech deterioration which is both acoustically robust and phonetically systematic. Both acoustic and magnetic resonance imaging measures were used to describe the speech disturbance in two neurologically distinct groups of cerebellar ataxia: Friedreich's ataxia and olivo-ponto cerebellar ataxia. The speaking task was designed to elicit six different prosodic conditions and four prosodic contrasts. All subjects read the same syllable embedded in a sentence, under six different prosodic conditions. Pair-wise comparisons derived from the six conditions were used to describe (1) final lengthening, (2) phrasal accent, (3) nuclear accent and (4) syllable reduction. An estimate of speech deterioration as determined by individual and normal subects' acoustic values of syllable duration, formant and fundamental frequencies was used in correlation analyses with magnetic resonance imaging ratings.

1pSC9. Effect of unilateral versus bilateral electrostimulation in subthalamic nucleus on speech in Parkinsons disease. Emily Wang (Dept. of Commun. Disord. and Sci., Rush Univ. Med. Ctr., 1653 W. Congress Pkwy., Ste. 203 SENN, Chicago, IL 60612), Leo Verhagen Metman (Rush Univ. Med. Ctr., Chicago, IL), Roy Bakay, Jean Arzbaecher (Chicago Inst. of Neurosurgery and Neuroresearch, Chicago, IL), and Bryan Bernard (Rush Univ. Med. Ctr., Chicago, IL)

Previously, it was found that 16 right-handed patients with idiopathic Parkinsons disease who underwent unilateral implantation of deep brain stimulator in subthalamic nucleus (STN) showed significant improvement in their nonspeech motor functions. Eight of the 16 patients had stimulator in the left STN and eight in the right STN. In contrast, their speech function showed very mild improvement that was limited to the respiratory/phonotory subsystems. Further, there seemed a trend that the patients with right STN stimulation did better than those with left STN stimulation. It was speculated that the difference might be due to a micro lesion caused by the surgical procedure to the corticobulbar fibers run in the left internal capsule. This paper reports speech changes associated with bilateral DBS in STN in four of the 16 subjects who elected to have deep brain stimulator implanted in STN on the opposite side of the brain at a later time. Results show negative changes in speech after bilateral DBS in STN. The changes were not limited to the micro lesion effect due to the surgery itself, but also related to the active stimulation on the dominant hemisphere for speech processing. [Work supported by NIH.]

1pSC10. The effect of noise on the perception of phonetic features in acoustic simulations of cochlear implant speech. Benjamin Munson, Peggy B. Nelson, and Jill E. Muecke (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Previous research [Munson and Nelson, J. Acoust. Soc. Am. **114**, 2360 (2003)] found that adults with cochlear implants (CIs) perceive synthetic /ra/-/la/, /wa/-/ja/, and *say-stay* continua less accurately and less categorically than listeners with normal-hearing sensitivity (NH) in noise. Considerable variability was found among the CI listeners: some demonstrated performance comparable to that of NH listeners, while others showed considerably poorer performance. One potential reason for the variability in performance is the number of distinct spectral channels available to the CI listeners. Friesen *et al.* [J. Acoust. Soc. Am. **110**, 1150–1163 (2001)], for example, demonstrated that better-performing implant users showed improvements in consonant and vowel recognition when the number of signal spectral channels was increased, with asymptotic performance at seven channels. Poorer-performing users, however, showed no improvement in scores when the number of channels increased beyond four. To examine this hypothesis, NH listeners participated in an experiment in which they listened to 4- and 8-band acoustic simulations of synthetic continua in quiet and in steady speech-shaped noise at a $+10$ dB SNR. Analyses will focus on the effect of number of channels on identification accuracy and within-group variability.

1pSC11. Speech perception in gated noise: the effects of spectral/temporal resolution and auditory integration. Su-Hyun Jin and Peggy B. Nelson (Commun. Disord., Univ. of Minnesota, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

The current investigation examined differences between hearing impaired (HI) listeners and normal-hearing listeners (NH) in the amount of masking release (MR) for sentence recognition and syllable identification tasks when they listened in modulated noise. HI and NH listeners' performance was compared when the performance levels of the two groups were equal in steady noise and in quiet. The relationships between the amount of MR to hearing threshold and suprathreshold abilities of forward masking, auditory filter bandwidth, and auditory stream segregation/integration were also investigated. To compensate for reduced hearing sensitivity for HI listeners, the spectrum levels of both speech and noise were adjusted based on the individual hearing loss. There was no significant performance difference between NH and HI groups in steady noise and in quiet. However, the amount of MR for sentences and for CV syllables was significantly reduced for HI listeners. For sentence recognition, the amount of MR seemed to be more related to hearing sensitivity for low-to-mid frequencies and the characteristics of auditory filters. Performance for gated sentence recognition was also strongly correlated with sentence recognition in gated noise. In contrast, forward masking thresholds appear to be the main contributor to the amount of MR for syllable recognition.

1pSC12. Vowel formant movement and duration perceived through noise vocoders and cochlear implants. Paul Iverson, Bronwen G. Evans, and Charlotte A. Smith (Dept. of Phonet. and Linguist., Univ. College London, 4 Stephenson Way, London NW1 2HE, UK, paul@phonetics.ucl.ac.uk)

Formant movement and duration have been increasingly shown to be important cues for vowel recognition by normal-hearing adults; individuals enhance formant movement and duration contrasts when speaking clearly, and vowel recognition accuracy declines when these differences are reduced in signal-processed or synthesized speech. This study investigated how these cues contribute to vowel recognition by cochlear implant users and normal-hearing individuals listening to noise-vocoded speech. Individuals were tested on vowel recognition using stimuli with and without formant movement and duration cues, and performed a goodness-rating task that found best exemplars for vowels that had formant movement and duration variation. The results suggested that duration was an

important cue to recognition for all listeners; removing duration reduced recognition accuracy by 6–19 percentage points in all conditions. Removing formant movement cues reduced the recognition of diphthongs to chance levels, but had little effect on the recognition of monophthongs. Goodness-rating data demonstrated that all listeners preferred systematic patterns of formant movement and duration variation, as occurs in natural speech. The results suggest that cochlear implant users and normal-hearing individuals give similar cue weightings to vowel formant movement and duration, but that formant movement may have less importance under these conditions than in previous studies.

1pSC13. Aided speech recognition in single-talker competition by elderly hearing-impaired listeners. Maureen Coughlin and Larry Humes (Speech and Hearing Dept., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, mcoughli@indiana.edu)

This study examined the speech-identification performance in one-talker interference conditions that increased in complexity while audibility was ensured over a wide bandwidth (200–4000 Hz). Factorial combinations of three independent variables were used to vary the amount of informational masking. These variables were: (1) competition playback direction (forward or reverse); (2) gender match between target and competition talkers (same or different); and (3) target talker uncertainty (one of three possible talkers from trial to trial). Four groups of listeners, two elderly hearing-impaired groups differing in age (65–74 and 75–84 years) and two young normal-hearing groups, were tested. One of the groups of young normal-hearing listeners was tested under acoustically equivalent test conditions and one was tested under perceptually equivalent test conditions. The effect of each independent variable on speech-identification performance and informational masking was generally consistent with expectations. Group differences in the observed informational masking were most pronounced for the oldest group of hearing-impaired listeners. The eight measures of speech-identification performance were found to be strongly correlated with one another, and individual differences in speech understanding performance among the elderly were found to be associated with age and level of education. [Work supported, in part, by NIA.]

1pSC14. Vowel acquisition by prelingually deaf children with cochlear implants. Marie-Eve Bouchard (Cognit. Neurosci. Ctr., Univ. of Quebec at Montreal, Montreal, QC, Canada), Marie-Thérèse Le Normand (Hopital de la Salpêtrière, Paris, France), Lucie Ménard, Marilyne Goud, and Henri Cohen (Univ. of Quebec at Montreal, Montreal, QC, Canada)

Phonetic transcriptions (study 1) and acoustic analysis (study 2) were used to clarify the nature and rhythm of vowel acquisition following the cochlear implantation of prelingually deaf children. In the first study, seven children were divided according to their degree of hearing loss (DHL): DHL I: 90–100 dB of hearing loss, 1 children; DHL II: 100–110 dB, 3 children; and DHL III: over 110 dB, 3 children. Spontaneous speech productions were recorded and videotaped 6 and 12 months postsurgery and vowel inventories were obtained by listing all vowels that occurred at least twice in the child's repertoire at the time of recording. Results showed that degree of hearing loss and age at implantation have a significant impact on vowel acquisition. Indeed, DHL I and II children demonstrated more diversified as well as more typical pattern of acquisition. In the second study, the values of the first and second formants were extracted. The results suggest evolving use of the acoustic space, reflecting the use of auditory feedback to produce the three phonological features exploited to contrast French vowels (height, place of articulation, and rounding). The possible influence of visual feedback before cochlear implant is discussed.

1pSC15. Optimizing cochlear implant frequency boundary tables for vowel perception: A computer simulation. Marios S. Fourakis (Speech & Hearing Sci., Ohio State Univ., 1070 Carmack Rd., Columbus, OH 43210), John W. Hawks (Kent State Univ., Kent, OH 44242, jhawks@kent.edu), and Amy Schwager (Ohio State Univ., Columbus, OH 43210)

For cochlear implants, the assignment of frequency bands to electrodes is a variable parameter that determines what region of the acoustic spectrum will be represented by each electrode's output. Technology will soon allow for considerable flexibility in programming this parameter. In a first attempt to optimize these assignments for vowel perception, a computer program was written to evaluate different assignment schemes for categorization accuracy based strictly on the frequency values of the first two or three formants. Databases [J. Hillenbrand *et al.*, J. Acoust. Soc. Am. **97**, 3099–3111 (1995)] of formant measurements from American English vowels as uttered by men, women, and children were used. For this simulation, it was assumed that each formant frequency was associated with only the frequency band its center frequency fell within. Each pattern of frequency bands was assigned a vowel category identification based on the plurality of tokens whose intended identification category fell within that pattern. A range of frequency scaling schemes for 19 and 20 electrode arrays was evaluated, with the best of these fine tuned for minimum error. The results indicate that manufacturer's default assignments categorize reasonably well, but a bark-scaled scheme yielded the best unmodified classifications.

1pSC16. An across-frequency deficit in hearing-impaired listeners is supported by acoustic correlation. Eric W. Healy, Anand Kannabiran (Dept. of Commun. Sci. and Disord., Univ. of South Carolina, Columbia, SC 29208, ewh@sc.edu), and Sid P. Bacon (Arizona State Univ., Tempe, AZ 85287-0102)

It has been recently suggested that listeners having a sensorineural hearing impairment (HI) may possess a deficit in their ability to integrate speech information across different frequencies. When presented with a task that required across-frequency integration of speech patterns and controlled for other known deficits, HI listeners performed more poorly than their normal-hearing (NH) counterparts [C. W. Turner *et al.*, J. Speech Lang. Hear. Res. **42**, 773–784 (1999); E. W. Healy and S. P. Bacon, J. Speech Lang. Hear. Res. **45**, 1262–1275 (2002)]. The latter study also showed that HI performance fell more steeply when increasing amounts of temporal asynchrony were introduced to the pair of widely separated patterns. In the current study, the correlations between the fluctuating envelopes of the acoustic stimuli were calculated, both when the patterns were aligned and also at various between-band asynchronies. It was found that the rate at which acoustic correlation fell as a function of asynchrony closely matched the rate at which intelligibility fell for the NH listeners. However, the intelligibility scores produced by the HI listeners fell more steeply than the acoustic analysis would suggest. Thus, these data provide additional support for the existence of an across-frequency deficit in HI listeners. [Work supported by NIH.]

1pSC17. Miller and Nicely's confusion data are predicted by Fletcher's AI. Jont B. Allen (Univ. of Illinois, Urbana, IL 61853, jba@auditorymodels.org)

Starting from the confusion matrix data $P_{ij}(\mathrm{SNR})$ of Miller and Nicely, where i indexes the stimulus and j the response, the average performance intensity (PI) function, over the 16 consonants, may be defined as $\bar{P}(\mathrm{SNR})=\Sigma_{i=1}^{16}P_{ii}(\mathrm{SNR})$. These data are for five female talkers. The average speech power for five female talkers is known from Dunn and White. Thus, the articulation index (AI) may be computed at each SNR value (the maximum AI value is 0.6). This allows one to re−express $\bar{P}(\mathrm{SNR})$ as $\bar{P}(\mathrm{AI})$. When the resulting function is compared to Fletcher′s formula for the AI, $P(\mathrm{AI})=1\text{-}e_{\min}^{\mathrm{AI}}$, with $e_{\min}=0.015$, the agreement is nearly perfect. Thus, the Miller and Nicely average phone data may be modeled by AI theory. This result seems astounding, as it falls outside

normal realm of the AI, which was intended for a much larger mix of sounds, not 16 consonants and a fixed vowel. Most individual consonant PI functions obey the same Fletcher AI formula, but with different values of $e_{\min}$.

1pSC18. Grouping Miller–Nicely by linear vector space rotations. Suvrat Budhlakoti, Jont B. Allen, and Erik Larsen (Beckman Inst. for Adv. Sci. and Technol., Univ. of Illinois at Urbana–Champaign, 405 N. Mathews Ave., Urbana, IL 61801)

Human speech recognition has been studied using response to CV speech stimuli. Miller and Nicely (1955) studied such data in the form of confusion matrices to obtain insight into the psychological structure of the phone in noise. Here, the confusion matrices are modeled as phone coordinates in a high dimensional perceptual vector space. The model generalizes to an eigenvalue decomposition (EVD) [Allen (2004)]. This is followed by agglomerative hierarchical clustering of the transformed data, and an automated process is used to identify the main clusters. The resulting EVD clustering is very similar to other Miller–Nicely groupings, based on both production and MDS derived features, but is more model based. It was found that there is a gradual and highly consistent change in the clustering of sounds, independent of cluster size and configuration. By examining the change in similarity between various speech sounds, it is hoped that perceptual features may be uniquely identified.

1pSC19. Development of context dependent sequential K-nearest neighbor classifier for usable speech classification. J. K. Shah, A. N. Iyer, B. Y. Smolenski, and R. E. Yantorno (Speech Processing Lab., College of Eng., Temple Univ., Philadelphia, PA 19122, shah@temple.edu)

The accuracy of speech processing applications degrades when operating in co-channel environment. Co-channel speech occurs when more than one person is talking at the same time. The idea of usable speech segmentation is to identify and extract those portions of co-channel speech that are minimally degraded but still useful for speech processing applications (such as speaker identification or speech recognition) which do not work in co-channel environments. Usable speech measures are features that are extracted from the co-channel signal to distinguish between usable and unusable speech. Several usable speech extraction methods have recently been developed based on a single feature of the speech signal being considered. In this paper, however, a new usable speech extraction technique, which sequentially and contextually selects several features of the given signal using the K-nearest neighbor classifier, is being investigated. This new approach considers periodicity and structure based features simultaneously in order to achieve the maximum classification rate, and by observing all the incoming frames, avoids the problem of deciding the amount of data needed to make accurate decisions. A 100% accuracy can be achieved in speech processing applications by using this extracted usable speech segment.

1pSC20. Usable speech processing: A novel approach to processing speech in degraded environments. Brett Y. Smolenski (Temple Univ., 12th and Norris, Philadelphia, PA 19122, bsmolens@temple.edu)

One of the main challenges still plaguing speech processing applications is enabling them to work in operational environments where interference and noise abound. The traditional approach has been to use some form of adaptive filtering operation. However, since the speech is nonstationary, it is possible to extract segments from the speech signal that have a large segmental signal-to-noise ratio (SNR) even when the overall SNR is very low. Such high SNR segments frequently occur during voiced speech, and experiments have shown that, using a speaker identification system, these high SNR segments can be correctly identified even when the entire utterance cannot. However, the segmental SNR is not normally known *a priori*. In this research, a statistical model is first developed for the segmental SNR values for several commonly occurring environments. It is then shown that a modified sinusoidal model and the Teager energy operator can be combined using a context dependent nonlinear regression technique to obtain a low variance estimate of the segmental SNR values provided the frame size is larger than 30 samples.

1pSC21. Unsupervised learning of broad phonetic classes with a statistical mixture model. Ying Lin (Phonet. Lab., Linguist. Dept., UCLA, Los Angeles, CA 90095-1543, yinglin@ucla.edu)

Unsupervised learning of broad phonetic classes by infants was simulated using a statistical mixture model. A mixture model assumes that data are generated by a certain number of different sources—in this case, broad phonetic classes. With the phonetic labels removed, hand-transcribed segments from the TIMIT database were used in model-based clustering to obtain data-driven classes. Simple hidden Markov models were chosen to be the components of the mixture, with mel-cepstral coefficients as the front end. The mixture model was trained using an expectation-maximization-like algorithm. The EM-like algorithm was initialized by a K-means procedure and then applied to estimate the parameters of the mixture model after iteratively partitioning the clusters. The results of running this algorithm on the TIMIT segments suggested that the partitions may be interpreted as gradient acoustic features, and that to some degree the resulting clusters correspond to knowledge-based phonetic classes. Although such correspondences are rather rough, a careful examination of the clusters showed that the class membership of some sounds is highly dependent on their phonetic contexts. Thus, the clusters may reflect the preliminary phonological categories formed during language learning in early childhood.

1pSC22. The influence of phonological priming on variability in articulation. Molly E. Babel and Benjamin Munson (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

Previous research [Sevald and Dell, Cognition **53**, 91–127 (1994)] has found that reiterant sequences of CVC words are produced more quickly when the prime word and target word share VC sequences (i.e., sequences like *sit sick*) than when they are identical (sequences like *sick sick*). Even slower production rates are found when primes and targets share a CV sequence (sequences like *kick sick*). These data have been used to support a model of speech production in which lexical items and their constituent phonemes are activated sequentially. The current experiment investigated whether phonological priming also influences variability in the acoustic characteristics of words. Specifically, we examined whether greater variability in the acoustic characteristics of target words was noted in the CV-related prime context than in the identical-prime context, and whether less variability was noted in the VC-related context. Thirty adult subjects with typical speech, language, and hearing ability produced reiterant two-word sequences that varied in their phonological similarity. The duration, first, and second formant frequencies of the target-words' vowels were measured. Preliminary analyses indicate that phonological priming does not have a systematic effect on variability in these acoustic parameters.

1pSC23. Noise robust speech recognition with support vector learning algorithms. Hassan H. Namarvar and Theodore W. Berger (Dept. of Biomed. Eng., Univ. of Southern California, OHE-500, Los Angeles, CA 90089-1451, heidarin@usc.edu)

We propose a new noise robust speech recognition system using time-frequency domain analysis and radial basis function (RBF) support vector machines (SVM). Here, we ignore the effects of correlative and nonstationary noise and only focus on continuous additive Gaussian white noise. We then develop an isolated digit/command recognizer and compare its performance to two other systems, in which the SVM classifier has been replaced by multilayer perceptron (MLP) and RBF neural networks. All systems are trained under the low signal-to-noise ratio (SNR) condition.

We obtained the best correct classification rate of 83% and 52% for digit recognition on the TI-46 corpus for the SVM and MLP systems, respectively under the SNR=0 (dB), while we could not train the RBF network for the same dataset. The newly developed speech recognition system seems to be noise robust for medium size speech recognition problems under continuous, stationary background noise. However, it is still required to test the system under realistic noisy environment to observe whether the system keeps its adaptability and robustness under such conditions. [Work supported in part by grants from DARPA CBS, NASA, and ONR.]

1pSC24. Intelligibility of an ASR-controlled synthetic talking face. Catherine Siciliano, Geoff Williams, Andrew Faulkner (Dept. of Phonet. and Linguist., Univ. College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK), and Giampiero Salvi (KTH, 10044 Stockholm, Sweden)

The goal of the SYNFACE project is to develop a multilingual synthetic talking face, driven by an automatic speech recognizer (ASR), to assist hearing-impaired people with telephone communication. Previous multilingual experiments with the synthetic face have shown that time-aligned synthesized visual face movements can enhance speech intelligibility in normal-hearing and hearing-impaired users [C. Siciliano *et al.*, Proc. Int. Cong. Phon. Sci. (2003)]. Similar experiments are in progress to examine whether the synthetic face remains intelligible when driven by ASR output. The recognizer produces phonetic output in real time, in order to drive the synthetic face while maintaining normal dialogue turn-taking. Acoustic modeling was performed with a neural network, while an HMM was used for decoding. The recognizer was trained on the SpeechDAT telephone speech corpus. Preliminary results suggest that the currently achieved recognition performance of around 60% frames correct limits the usefulness of the synthetic face movements. This is particularly true for consonants, where correct place of articulation is especially important for visual intelligibility. Errors in the alignment of phone boundaries representative of those arising in the ASR output were also shown to decrease audio-visual intelligibility. [Work supported by the EU IST Project 2001-33327.]

1pSC25. A novel dynamic acoustical model for speaker verification. Gongjun Li and Carol Espy-Wilson (Dept. of Elec. and Computer Eng., Univ. of Maryland at College Park, College Park, MD 20742, gongjun@glue.umd.edu)

In speaker verification, the conventional acoustical models (hidden Markov model and vector quantization) are not able to capture a speaker's dynamic characteristics. In this paper we describe a novel dynamic acoustical model. The training data are viewed as a concatenation of many speech-pattern samples, and the pattern matching involves a comparison of the pattern samples and the test speech. To reduce the amount of computation, a tree is generated to index the entrance to pattern samples using an expectation and maximization (EM) approach, and leaves in the tree are employed to quantize the feature vectors in the training data. The obtained leaf-number sequences are exploited in pattern matching as a temporal model. We use a DTW scheme and a GMM scheme to match the training data and the test speech. Experimental results on NIST'98 speaker recognition evaluation data show that the accuracy of speaker verification on 3- and 10-s test speech is raised from 71.1% and 75.2% for a baseline GMM-based system to 80.0% and 82.1% for the dynamic acoustical model, respectively. Furthermore, some pattern samples in the training data are correctly tracked by the test speech.

1pSC26. Automatic detection of the features [high] and [low] in a landmark-based model of speech perception. Janet Slifka (Res. Lab. of Electron., MIT, 36-587, 50 Vassar St., Cambridge, MA 02139)

This research is part of a landmark-based approach to modeling speech perception in which sound segments are assumed to be represented as bundles of binary distinctive features. In this model, probability estimates for feature values are derived from measurements of the acoustics in the vicinity of landmarks. The goal of the current project is to automatically detect the features [high] and [low] for vowel segments based on measurements from average spectra. A long-term and a short-term average spectrum are computed using all vowel regions in the utterance and are used to estimate speaker-specific parameters such as average $F0$ and average $F3$ (an indicator of vocal tract length). These parameters are used to estimate $F1$ using a peak-picking process on the average spectrum at each vowel-landmark. Preliminary results are derived from read connected speech for 738 vowels from 80 utterances (two male speakers, two female speakers). Speaker-independent logistic regression analysis using only average $F0$ and $F1$ determines the feature [high] with 73% accuracy and the feature [low] with 84% accuracy. Proposals are made for methods to use additional spectral detail to create a more robust estimate for vowels which show significant formant movement. [Work supported by NIH Grant No. DC02978.]

1pSC27. Spectral variability at the transition between successive phonemes. Sorin Dusan (Ctr. for Adv. Information Processing, Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854, sdusan@caip.rutgers.edu)

In an experimental study of identification of truncated Japanese syllables it was stated that a short speech interval (approximately 10 ms) that includes the position of maximum spectral transition between a consonant and a vowel carries the most important information for the perception of the consonant and the syllable [S. Furui, J. Acoust. Soc. Am. **80**, 1016–1025 (1986)]. A reduced spectral variability at the transition position could partially explain the increase of information at this position. The current study investigates whether there is a decrease in spectral variability at the transition position between successive phonemes compared with the spectral variability at the phoneme centers. The training part of the TIMIT acoustic-phonetic database, containing sentences in English from 462 American speakers, is used to build 2471 diphone models, based on the 61 symbols used in the database for phonetic transcription. The variability of the mel-frequency cepstral coefficients (MFCCs) is evaluated for various diphone models at the phoneme centers and at the phoneme transition position. Preliminary results suggest that spectral variability is not significantly lower at the phoneme transition positions than that at the phoneme centers in these diphone models.

1pSC28. Automatic speech recognizer based on the Spanish spoken in Valdivia, Chile. Maria L. Sanchez (Escuela de Ingenieria Acustica, Facultad de Ciencias de la Ingenieria, Universidad Austral de Chile, General Lagos 2086, Valdivia, Chile), Victor H. Poblete, and Jorge Sommerhoff (Universidad Austral de Chile, Valdivia, Chile)

The performance of an automatic speech recognizer is affected by training process (dependent on or independent of the speaker) and the size of the vocabulary. The language used in this study was the Spanish spoken in the city of Valdivia, Chile. A representative sample of 14 students and six professionals all natives of Valdivia (ten women and ten men) were used to complete the study. The sample ranged in age between 20 and 30 years old. Two systems were programmed based on the classical principles: digitalizing, end point detection, linear prediction coding, cepstral coefficients, dynamic time warping, and a final decision stage with a previous step of training: (i) one dependent speaker (15 words: five colors and ten numbers), (ii) one independent speaker (30 words: ten verbs, ten nouns, and ten adjectives). A simple didactical application, with options to choose colors, numbers and drawings of the verbs, nouns and adjectives, was designed to be used with a personal computer. In both programs, the

tests carried out showed a tendency towards errors in short words with monosyllables like ''flor,'' and ''sol.'' The best results were obtained in words with three syllables like ''disparar'' and ''mojado.'' [Work supported by Proyecto DID UACh N S-200278.]

1pSC29. Auditory analysis for speech recognition based on physiological models. Woojay Jeon and Biing-Hwang Juang (School of Elec. and Computer Eng., Georgia Inst. of Technol., Atlanta, GA 30332, wjjeon@ece.gatech.edu)

To address the limitations of traditional cepstrum or LPC based front-end processing methods for automatic speech recognition, more elaborate methods based on physiological models of the human auditory system may be used to achieve more robust speech recognition in adverse environments. For this purpose, a modified version of a model of the primary auditory cortex featuring a three dimensional mapping of auditory spectra [Wang and Shamma, IEEE Trans. Speech Audio Process. **3**, 382–395 (1995)] is adopted and investigated for its use as an improved front-end processing method. The study is conducted in two ways: first, by relating the model's redundant representation to traditional spectral representations and showing that the former not only encompasses information provided by the latter, but also reveals more relevant information that makes it superior in describing the identifying features of speech signals; and second, by observing the statistical features of the representation for various classes of sound to show how different identifying features manifest themselves as specific patterns on the cortical map, thereby becoming a place-coded data set on which detection theory could be applied to simulate auditory perception and cognition.

1pSC30. Measures of voiced frication for automatic classification. Philip J. B. Jackson (Ctr. for Vision, Speech and Signal Processing, Univ. of Surrey, Guildford GU2 7XH, UK), Luis M. T. Jesus (Universidade de Aveiro, 3810-193 Aveiro, Portugal), Christine H. Shadle (Univ. of Southampton, Southampton SO17 1BJ, UK), and Jonathan Pincas (Univ. of Surrey, Guildford GU2 7XH, UK)

As an approach to understanding the characteristics of the acoustic sources in voiced fricatives, it seems apt to draw on knowledge of vowels and voiceless fricatives, which have been relatively well studied. However, the presence of both phonation and frication in these mixed-source sounds offers the possibility of mutual interaction effects, with variations across place of articulation. This paper examines the acoustic and articulatory consequences of these interactions and explores automatic techniques for finding parametric and statistical descriptions of these phenomena. A reliable and consistent set of such acoustic cues could be used for phonetic classification or speech recognition. Following work on devoicing of European Portuguese voiced fricatives [Jesus and Shadle, in Mamede *et al.* (eds.) (Springer-Verlag, Berlin, 2003), pp. 1–8]. and the modulating effect of voicing on frication [Jackson and Shadle, J. Acoust. Soc. Am. **108**, 1421–1434 (2000)], the present study focuses on three types of information: (*i*) sequences and durations of acoustic events in VC transitions, (*ii*) temporal, spectral and modulation measures from the periodic and aperiodic components of the acoustic signal, and (*iii*) voicing activity derived from simultaneous EGG data. Analysis of interactions observed in British/American English and European Portuguese speech corpora will be compared, and the principal findings discussed.

1pSC31. Regularized reestimation of stochastic duration models for phone-classification. Martin J. Russell (Electron., Elec. and Computer Eng., Univ. of Birmingham, Edgbaston, Birmingham B15 2TT, UK) and Philip J. B. Jackson (Univ. of Surrey, Guildford GU2 7XH, UK)

Recent research has compared the performance of various distributions (uniform, boxcar, exponential, gamma, discrete) for modeling segment (state) durations in hidden semi-Markov models used for phone classification on the TIMIT database. These experiments have shown that a gamma distribution is more appropriate than exponential (which is implicit in first-order Markov models), and achieved a 3% relative reduction in phone-classification errors [Jackson, Proc. ICPhS, pp. 1349–1352 (2003)]. The parameters of these duration distributions were estimated once for each model from initial statistics of state occupation (offline), and remained unchanged during subsequent iterations of training. The present work investigates the effect of reestimating the duration models in training (online) with respect to the phone-classification scores. First, tests were conducted on duration models reestimated directly from statistics gathered in the previous iteration of training. It was found that the boxcar and gamma models were unstable, meanwhile the performance of the other models also tended to degrade. Secondary tests, using a scheme of annealed regularization, demonstrated that the losses could be recouped and a further 1% improvement was obtained. The results from this pilot study imply that similar gains in recognition accuracy deserve investigation, along with further optimization of the duration model reestimation procedure.

1pSC32. The influence of semantic information on the acoustics of speech in noise. Rupal Patel, Mariam Syeda, and Aviva Krauthammer (Dept. of Speech Lang. Pathol. and Audiol., Northeastern Univ., 360 Huntington Ave., 102 FR, Boston, MA 02115, r.patel@neu.edu)

While there is a significant body of work on how people modify their speech patterns in the presence of noise, the role of semantic information on these acoustic modifications is not well understood. This study examined whether adult speakers of English differentially modify semantically salient versus nonsalient words within a sentence in the presence of noise. Participants were asked to produce a set of 20 sentences [from the Speech Perception in Noise List, Kalikow *et al.* (1977)], in each of five noise conditions; quiet, 60 and 90 dB SPL multispeaker conversation, and 60 and 90 dB SPL street noise. Five random repetitions per sentence were requested for each noise condition. The following acoustic cues were extracted for each word within an utterance: duration, average intensity, peak intensity, average fundamental frequency, and peak fundamental frequency. The ratios of these measures were then compared for semantically salient versus nonsalient words. While we expect to see an overall change in all of these prosodic cues for speech produced in noise (Lombard effect), semantically salient words may exhibit a greater ratio of change in some or all features.

1pSC33. The relation between semantics and lexical properties in spoken word production. Patricia Amico, Jan Charles-Luce, and Elizabeth McEldowney (Dept. of Commun. Disord. and Sci., 122 Cary Hall, Univ. at Buffalo, Buffalo, NY 14214)

Previous research has demonstrated facilitation in speech production when multiple levels are activated. In the present study a time course of facilitation and the contribution of various lexical and semantic properties in facilitating spoken word production was investigated. An experiment was conducted that examined the effects of semantics and lexical properties on acoustic-phonetic duration on spoken word production. Specifically, the primary interest was how duration changed as a function of semantic context and its interaction with the frequency and similarity neighborhood of CVC words. The semantic contexts and targets were presented using a visual naming paradigm. Target words were presented either 100 or 1500 ms after the offset of a simultaneous presentation of a string of three primes. The three primes were either all semantically related to the target, all semantically unrelated to the target, or all nonlinguistic characters neutral to the target. Duration of the target stimuli and reaction times to onset of articulation were measured. The results will be discussed in terms of interactive activation. [Work supported by NIH NIDCD Grant R01 0265801.]

1pSC34. Adaptation to structural modifications of the human vocal tract during speech: Electropalatographic measures. Wendi A. Aasland, Shari R. Baum (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, wendi.aasland@mail.mcgill.ca), and David H. McFarland (Université de Montréal, Montréal, QC, Canada)

Structural modifications to the vocal tract force speakers to alter their previously learned articulatory patterns in order to produce perceptually adequate speech. Previous research has shown that acoustic output in the production of alveolar consonants changes during adaptation to structural alterations of the palate, but to date, little is known regarding exactly how these changes result kinematically. The present study examines the adjustments made to tongue–palate contact patterns, measured using electropalatography (EPG), during adaptation to a palatal perturbation for the fricative [s]. Productions of the nonsense word [asa] were elicited in nine subjects at five time intervals, 15 min apart, while speakers wore electropalates modified with a thicker-than-normal alveolar ridge. Between measurement intervals, speakers read [s]-laden passages to promote adaptation. Productions were also elicited with an unperturbed electropalate in place to characterize normal articulation. Electropalatographic analyses revealed a posterior shift in center of gravity of tongue–palate contact, alterations in the width of the medial groove necessary for [s] production, and increased variability in productions, which may reflect the instability of the new motor programs. Results are discussed in relation to the development of adaptive articulatory programs in speech motor control. [Work supported by NSERC and a FRSQ Bourse de Formation.]

1pSC35. TADA: An enhanced, portable Task Dynamics model in MATLAB. Hosung Nam, Louis Goldstein (Haskins Labs. & Yale Univ., 270 Crown St., New Haven, CT 06511, hosung.nam@yale.edu), Elliot Saltzman (Boston Univ., Boston, MA 02215), and Dani Byrd (USC Linguist. and Haskins Labs., Los Angeles, CA 90089-1693)

A portable computational system called TADA was developed for the Task Dynamic model of speech motor control [Saltzman and Munhall, Ecol. Psychol. **1**, 333–382 (1989)]. The model maps from a set of linguistic gestures, specified as activation functions with corresponding constriction goal parameters, to time functions for a set of model articulators. The original Task Dynamic code was ported to the (relatively) platform-independent MATLAB environment and includes a MATLAB version of the Haskins articulatory synthesizer, so that articulator motions computed by the Task Dynamic model can be used to generate sound. Gestural scores can now be edited graphically and the effects of gestural score changes on the models output evaluated. Other new features of the system include: (1) A graphical user interface that displays the input gestural scores, output time functions of constriction goal variables and articulators, and an animation of the resulting vocal-tract motion; (2) Integration of the Task Dynamic model with the prosodic clock-slowing, pi-gesture model of Byrd and Saltzman [J. Phonetics **31**, 149–180 (2003)]. This now allows prosody-driven slowing to be applied to the full set of active gestures and its effects to be evaluated perceptually. [Work supported by NIH.]

1pSC36. The role of coda consonants in triggering speech errors: An ultrasound study. Marianne Pouplier (Vocal Tract Visualization Lab, UMAB Dental School, BMS, 666 W. Baltimore St., Baltimore, MD 21201, mpoup001@umaryland.edu)

Recent speech-error research using articulatory data has shown that errors can often result in a phonologically illegal structure. These findings have been interpreted to mean that speech production is a fundamentally coordinative process. In speech errors, gestures can fall into grammatically illegal, albeit dynamically stable coordination modes. Shared gestural structures, such as identical consonants in the vicinity of the gestures affected by error, set up the conditions under which dynamically stable coordination modes can come to dominate over grammatical coordination patterns (e.g., Pouplier, 2003). This approach finds support in the often-observed phenomenon that a shared final consonant between two words will increase the likelihood of the initial consonants interacting. The present study uses ultrasound data of speech errors to investigate in what domains other than the coda shared gestural structure can trigger errors. Stimuli without coda consonants are employed to investigate whether errors on the initial consonants can be triggered by a shared vowel only. It is further examined whether the shared structure triggering errors can reside within the initial consonants themselves, e.g., when the initial consonants are complex, multigestural constellations that overlap in some aspects of their gestural composition. [Work supported by NIH RO1-DC01758.]

1pSC37. Motor learning of volitional nonspeech oral movements: Intraoral pressure and articulatory kinematics. Susan Shaiman, Malcolm R. McNeil (Dept. of Commun. Sci. & Disord., Univ. of Pittsburgh, 4033 Forbes Tower, Pittsburgh, PA 15260, shaiman@csd.pitt.edu), and Neil J. Szuminsky (Necessity Consulting, Mars, PA 16046)

This study provides initial validation of a complex nonspeech task that will be used in future research for examining and comparing the neurophysiologic mechanisms of speech and volitional nonspeech oral movements. Motor learning of a complex sequence of lip, jaw, and tongue movements was explored as speakers produced one of three intraoral air-pressure targets during each bilabial closing gesture in the sequence. Motor learning was demonstrated by retention of pressure targets subsequent to acquisition. Transfer to other nonspeech tasks was also explored, as were changes in articulatory kinematics with learning. The current nonspeech task was constructed to parallel speech production, by controlling several commonly observed physiologic characteristics of speech (i.e., complex sequence of potentially overlapping articulatory segments; goal of intraoral pressure during bilabial closure). This construction of the nonspeech task, along with its demonstrated motor learning, greatly extends the ability to make valid comparisons between speech and nonspeech productions. [Work supported by CMRF-University of Pittsburgh.]

1pSC38. Sensorimotor adaptation to acoustic perturbations in vowel formants. Virgilio Villacorta, Joseph Perkell (Res. Lab. of Electron., MIT, 50 Vassar St., Cambridge, MA 02139, virgilio@mit.edu), and Frank Guenther (Boston Univ., Boston, MA 02215)

The goal of this research is to study the auditory component of feedback control in speech production. This experiment investigates auditory sensorimotor adaptation (SA) as it relates to speech production: the process by which speakers alter their speech production in order to compensate for perturbations of their normal auditory feedback. Specifically, the first formant frequency ($F1$) was shifted in the auditory feedback heard by naive adult subjects as they produced vowels in single syllable words. Initial results indicate that subjects demonstrate compensatory formant shifts in their speech. This compensation was also present after training when acoustic feedback was masked by noise. This suggests that internal models used in the control of speech movements can be constantly updated by auditory feedback. These results in voiced speech are consistent with results from Houde and Jordan [Science **279**, 1213–1216 (1998)], which demonstrated SA in whispered speech. A second study, currently underway, investigates perceptual discrimination of vowel stimuli differing in $F1$ frequency, using the same subjects as in the SA studies. Cross-subject relations between discrimination scores and extent of compensation will be presented and discussed. [Work supported by NIDCD Grant R01-DC01925.]

1pSC39. Statistical comparison of similarity tests applied to speech production data. H. Kollia (Dept. of Commun. Disord., William Paterson Univ., Wayne, NJ 07470), Jay Jorgenson, Rose Saint Fleur, and Kevin Foster (The City College of the City Univ. of New York, New York, NY 10031)

Statistical analysis of data variability in speech production research has traditionally been addressed with the assumption of normally distributed error terms. The correct and valid application of statistical procedure requires a thorough investigation of the assumptions that underlie the

methodology. In previous work [Kollia and Jorgenson, J. Acoust. Soc. Am. **102** (1997); **109** (2002)], it was shown that the error terms of speech production data in a linear regression can be modeled accurately using a quadratic probability distribution, rather than a normal distribution as is frequently assumed. The measurement used in the earlier Kollia–Jorgenson work involved the classical Kolmogorov–Smirnov statistical test. In the present work, the authors further explore the problem of analyzing the error terms coming from linear regression using a variety of known statistical tests, including, but not limited to chi-square, Kolmogorov–Smirnov, Anderson–Darling, Cramer–von Mises, skewness and kurtosis, and Durbin. Our study complements a similar study by Shapiro, Wilk, and Chen [J. Am. Stat. Assoc. (1968)]. [Partial support provided by PSC-CUNY and NSF to Jay Jorgenson.]

1pSC40. Spectral characteristics of speech with fixed jaw displacements. Nancy P. Solomon, Matthew J. Makashay (Army Audiol. and Speech Ctr., Walter Reed Army Med. Ctr., 6900 Georgia Ave. NW, Washington, DC 20307, nancy.solomon@na.amedd.army.mil), and Benjamin Munson (Univ. of Minnesota, Minneapolis, MN 55455)

During speech, movements of the mandible and the tongue are interdependent. For some research purposes, the mandible may be constrained to ensure independent tongue motion. To examine specific spectral characteristics of speech with different jaw positions, ten normal adults produced sentences with multiple instances of /t/, /s/, /ʃ/, /i/, /ai/, and /ɔi/. Talkers produced stimuli with the jaw free to vary, and while gently biting on 2- and 5-mm bite blocks unilaterally. Spectral moments of /s/ and /ʃ/ frication and /t/ bursts differed such that mean spectral energy decreased, and diffuseness and skewness increased with bite blocks. The specific size of the bite block had minimal effect on these results, which were most consistent for /s/. Formant analysis for the vocoids revealed lower $F2$ frequency in /i/ and at the end of the transition in /ai/ when bite blocks were used; $F2$ slope for diphthongs was not sensitive to differences in jaw position. Two potential explanations for these results involve the physical presence of the bite blocks in the lateral oral cavity, and the oromotor system's ability to compensate for fixed jaw displacements. [Work supported by NIDCD R03-DC06096.]

1pSC41. Segment sequencing and the jaw cycle. Melissa Redford (Dept. of Linguist., 1290 Univ. of Oregon, Eugene, OR 97403, redford@darkwing.uoregon.edu) and Paul van Donkelaar (Univ. of Oregon, Eugene, OR)

The open–close jaw cycle established in running speech often spans the articulation of several segments. Moreover, the extremes of the cycle are associated with opposing segment types, consonants and vowels, which form higher-level linguistic units, syllables, when grouped. The present study sought to determine how the syllable is related to the cycle. A partial relationship may imply that the jaw moves linearly from one segment to another. An isomorphic relationship that jaw movement is planned over the articulation of a group of segments, regardless of segment sequencing within the group. Five speakers repeated one of five syllable types (CV, VC, CVC, CCV, VCC) continuously for several seconds, producing an unbroken sequence of consonants and vowels. Jaw movement and acoustic data were recorded simultaneously and used to determine the temporal boundaries of the cycles and syllables, respectively. Cycle boundaries were defined by the points of maximum closure preceding and following an opening gesture, and syllable boundaries according to changes in frequency and amplitude corresponding to the onset of the initial and final segments in the syllable. A comparison of the relative boundary locations suggested an isomorphic relationship between cycles and syllables that interacted with segment type, syllable shape, and speaker.

1pSC42. A new 3D dynamical biomechanical tongue model. Jean-Michel Gerard, Pascal Perrier (ICP-INPG, 46 av. Felix Viallet, 38031 Grenoble, Cedex 1, France, gerard@icp.inpg.fr), Yohan Payan (Univ. Joseph Fourier, 38706 La Tronche, Cedex, France), and Reiner Wilhelms-Tricarico (MIT, Cambridge, MA)

A new dynamical biomechanical tongue model is being developed to study speech motor control. In spite of its computational complexity, a 3D representation was chosen in order to account for various contacts between tongue and external structures such as teeth, palate, and vocal tract walls. A fair representation of tongue muscle anatomy is provided, by designing the finite element mesh from the visible human data set (female subject). Model geometry was then matched to a human speaker, so that simulations can be quantitatively compared to experimental MRI data. A set of 11 muscles is modeled, whose role in speech gestures is well established. Each muscle is defined by a set of elements whose elastic properties change with muscle activation. Muscles forces are applied to the tongue model via macrofibers defined within the mesh by muscle specific sets of nodes. These forces are currently specified as step functions. Boundary conditions are set using zero-displacement nodes simulating attachments of tongue on bony structures. The nonlinear mechanical properties of tongue soft tissues are modeled using a hyperelastic material. Three-dimensional tongue deformations generated by each muscle, using FEM software ANSYS for computation, will be presented. Implications for speech motor control will be proposed.

1pSC43. Nonlinearity between acoustics and articulation in Hungarian transparent vowels. Stefan Benus, Karen D. Kirke, and Adamantios I. Gafos (Dept. of Linguist., New York Univ., 719 Broadway, 4th Fl., New York, NY 10003)

We present novel results from the acoustic and articulatory investigation of the production of the transparent vowels (TVs) /i/, /i:/, /e:/ in Hungarian (colon denotes length). The acoustic measurements of the front–back distinction (second formant, the difference of the first and second formants [Ladefoged, 1993]) show that the effect of adjacent back vowels on the front quality of the TVs is only weakly significant. The articulatory measurements of the same data, however, show that adjacent back vowels cause highly significant retraction of the tongue body during the production of the front TVs. The significance of this finding lies in its relevance to the relationship between phonetics and phonology. Our results demonstrate that minor phonetic differences in articulation, impossible to access by traditional theory, correlate with full-fledged phonological alternation of suffix selection in Hungarian. Traditional phonological accounts predict no effect of continuous phonetic details on discrete phonological generalizations. This is supported in our acoustic data but contrasts with our articulatory findings. In the paper we propose a dynamic model where phonological transparency is directly related to nonlinearity between acoustics and articulation [Stevens, 1989; Wood, 1979]. [Work supported by NIH.]

Session 1pSP

Signal Processing in Acoustics: Distinguished Lecture on Communication Acoustics

Ning Xiang, Chair
Architecture, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180

Chair's Introduction—1:15

Invited Paper

1:20

1pSP1. Communication acoustics. Jens Blauert (Inst. of Commun. Acoust., Ruhr-Universitaet Bochum, D-44780 Bochum, Germany, jens.blauert@rub.de)

Those aspects of acoustics which concern the relations of acoustics to the information and communication technologies are now frequently called ''communication acoustics.'' After a short review of the history of this field, relevant results from recent research at the Institute of Communication Acoustics at the Ruhr-University of Bochum, Germany, will be reported. This work can be seen in light of the research areas of computational auditory scene analysjs (CASA) and auditory virtual environments (AVE)—both dealing with the parametric representation of auditory scenes. Recent application opportunities and future trends will be discussed. It will be argued that modern communication–acoustical systems—which are often only embedded components in more complex communication systems—require more and more built-in explicit knowledge. Among other things, the development of such components and systems calls for data and knowledge from the cognitive sciences. Today's programs for education in communication engineering and communication acoustics often do not take sufficient account of this trend.

Session 1pST

ASA Committee on Standards: Role of Standards in ASA

Paul D. Schomer, Chair
Schomer & Associates, Inc., 2117 Robert Drive, Champaign, Illinois 61821

Chair's Introduction—1:15

Invited Papers

1:20

1pST1. Acoustical standards in the Society: Historical perspective. Tony F. W. Embleton (80 Sheardown Dr., Nobleton, ON L0G 1N0, Canada), Paul D. Schomer (Schomer and Assoc., Inc., Champaign, IL 61821), and Susan B. Blaeser (Acoust. Society of America Standards Secretariat, Melville, NY 11747)

The Acoustical Society of America has been active in national standards since 1932 when it asked the American Standards Association to initiate a project to standardize acoustical measurements and terminology. The new committee was assigned to the Society and designated Z24. By 1942, Z24 had expanded to include vibration. By 1957 the activity had grown so much that Z24 was split into three committees: S1 Acoustics, S2 Mechanical Vibration and Shock, and S3 Bioacoustics. (S12 Noise was added later, in 1981). Until 1969, these committees were administered by the USA Standards Institute, the predecessor of the American National standards Institute (ANSI), and new standards were approved by the Executive Council of the Society. In 1971, the Society assumed the responsibility for developing the US position on international standards in acoustics being generated by TC108 of the International Organization for Standardization. During the early 1970s ANSI became the body to approve all national standards, and the Soceity became a standards developer. ASA established a Committee on Standards (ASACOS) in 1978 to guide the work of the S-Committees and the Standards Secretariat. In the past few years ASACOS has increasingly improved its links with the Society's other technical activities.

1:40

1pST2. Standards 101: The ASA Standards program. Paul Schomer (Schomer and Assoc., Inc., 2117 Robert Dr., Champaign, IL 61821)

ASA serves as a standards developer under the auspices of the American National Standards Institute (ANSI). The Standards Program is organized through four technical committees (S1, S2, S3, and S12) and one administrative committee (ASACOS). S1 deals with physical acoustics, S2 deals with shock and vibration, S3 deals with physiological and psychological acoustics and S12 deals with noise. ASACOS is the ASA Committee on Standards. The program has three primary tasks: (1) development of national standards (ANSI Standards), (2) national adoption of international standards (ANSI NAIS Standards), (3) providing the USA input to the development of international standards (ISO and IEC Standards). At every level the main work is accomplished in Working Groups (WG) that are staffed by hundreds of volunteers, mainly ASA members from its various technical committees such as Noise, Physical Acoustics, Architectural Acoustics, Physiological and Psychological Acoustics, etc. Overall, the Standards Program involves more ASA members than does any other single function of the society except meetings. It is the biggest outreach function of ASA affecting the health, welfare, and economic well-being of large sectors of society. It is a main way the ASA diffuses the knowledge of acoustics and its practical application, perhaps the main way.

2:00

1pST3. The roles of major international standards organizations. Steven Cornish (Director of Intl. Policy, ANSI, 25 W. 43rd St., New York, NY 10036)

This review paper describes and analyzes the roles of the major international standards organizations, the International Organization for Standardization (ISO) and the International Electro-technical Commission (IEC). In particular, it analyzes the relationships of ISO and IEC with ANSI, CEN (including the ISO/CEN Vienna Agreement), and CENELEC. This paper also addresses recent developments to ensure the global relevance of ISO and IEC standards. Lastly, this paper addresses the myths, perceptions, and realities of European participation and potential dominance in ISO.

2:20

1pST4. The role of standards in noise. Richard J. Peppin (Scantek, Inc., 7060 Oakland Mills Rd. #L, Columbia, MD 21046, peppinr@asme.org)

This paper discusses how the noise community uses standards. The noise community consists of people working in the control and abatement of noise sources: produced by, or isolated as a result of, mechanical, information, transportation, vehicle, highway, industrial, and building systems. For the purposes of this paper, standards are described as ''agreed upon ways of doing things.'' Many of the standards available, their organizations, and their interrelationships will be addressed. The presentation will show how the various disciplines are concerned with the different groupings of standards and why, in many instances, it is impossible to do work without them.

2:40

1pST5. Role of standards in ASA: Animal bioacoustics. Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com)

Animal bioacoustics (AB) is a participant in ASACOS committee S3, Bioacoustics, and has one working group S3/WG90. However, standards that could be written for animals cut across committee boundaries, from development of terminology (S1), to specification of audiometric methods (S2), to hearing conservation programs (S12). From a biologist's perspective, there will be a fundamental difference between developing AB standards and those that have been published during the 75-year history of ASA—instead of focusing on one species, *Homo sapiens*, AB standards will be applied to many, even if data are available for only a few. Given the intensive research dedicated to establishing the existing standards for *H. sapiens*, the task ahead appears daunting. How should standards for thousands of species be specified? How should the standards process include a measure critical for wild animals, population sustainability? Writers of AB standards must apply a powerful conceptual tool, the comparative approach, and must design standards to incorporate new data quickly. Funding sources must recognize the need for these approaches. If they do, the standards developed will be marketable; it is also likely that important new perspectives on human bioacoustics will emerge.

3:00–3:10 Break

3:10

1pST6. The application of standard definitions of sound to the fields of underwater acoustics and acoustical oceanography. William M. Carey (College of Eng., Boston Univ., Boston, MA 02215, kerrygrp@ctol.net)

Recent societal concerns have focused attention on the use of sound as a probe to investigate the oceans and its use in naval sonar applications. The concern is the impact the use of sound may have on marine mammals and fishes. The focus has changed the fields of acoustical oceanography (AO) and underwater acoustics (UW) because of the requirement to communicate between disciplines. Multiple National Research Council publications, Dept. of Navy reports, and several monographs have been written on this subject, and each reveals the importance as well as the misapplication of ASA standards. The ANSI-ASA standards are comprehensive, however not widely applied. The clear definition of standards and recommendations of their use is needed for both scientists and government agencies. Traditionally the U.S. Navy has been responsible for UW standards and calibration; the ANSI-ASA standards

have been essential. However, recent changes in the Navy and its laboratory structure may necessitate a more formal recognition of ANSI-ASA standards and perhaps incorporation of UW-AO in the Bureau of Standards. A separate standard for acoustical terminology, reference levels, and notation used in the UW-AO is required. Since the problem is global, a standard should be compatible and cross referenced with the International Standard (CEI/IEC 27-3).

3:30

1pST7. ISO/IEC 17025 laboratory accreditation of NRC Acoustical Standards Program. George S. K. Wong, Lixue Wu, Peter Hanes, and Won-Suk Ohm (Inst. for Natl. Measurement Standards, Natl. Res. Council, Ottawa, ON K1A 0R6, Canada, george.wong@nrc-cnrc.gc.ca)

Experience gained during the external accreditation of the Acoustical Standards Program at the Institute for National Measurement Standards of the National Research Council is discussed. Some highlights include the preparation of documents for calibration procedures, control documents with attention to reducing future paper work and the need to maintain documentation or paper trails to satisfy the external assessors. General recommendations will be given for laboratories that are contemplating an external audit in accordance to the requirements of ISO/IEC 17025.

3:50

1pST8. Acoustical standards in engineering acoustics. Mahlon D. Burkhard (239 Stephen Ln., Charles Town, WV 25414)

The Engineering Acoustics Technical Committee is concerned with the evolution and improvement of acoustical techniques and apparatus, and with the promotion of new applications of acoustics. As cited in the Membership Directory and Handbook (2002), the interest areas include transducers and arrays; underwater acoustic systems; acoustical instrumentation and monitoring; applied sonics, promotion of useful effects, information gathering and transmission; audio engineering; acoustic holography and acoustic imaging; acoustic signal processing (equipment and techniques); and ultrasound and infrasound. Evident connections between engineering and standards are needs for calibration, consistent terminology, uniform presentation of data, reference levels, or design targets for product development. Thus for the acoustical engineer standards are both a tool for practices, for communication, and for comparison of his efforts with those of others. Development of many standards depends on knowledge of the way products are put together for the market place and acoustical engineers provide important input to the development of standards. Acoustical engineers and members of the Engineering Acoustics arm of the Society both benefit from and contribute to the Acoustical Standards of the Acoustical Society.

MONDAY EVENING, 24 MAY 2004 NEW YORK BALLROOM A, 7:00 TO 9:00 P.M.

Session 1eID

Interdisciplinary: Tutorial Lecture: Listening to the Acoustics in Concert Halls

Patricia K. Kuhl, Chair
University of Washington, UW Box 357920, Seattle, Washington 98195

Chair's Introduction—7:00

Invited Paper

7:05

1eID1. Listening to the acoustics in concert halls. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755) and David Griesinger (Lexicon, Bedford, MA 01730-1441)

How does acoustics affect the symphonic music performed in a concert hall? The lecture begins with an illustrated discussion of the architectural features that influence the acoustics. Boston Symphony Hall, which was built in 1900 when only one facet of architectural design was known, now rates as one of the world's great halls. How this occurred will be presented. Music is composed with some acoustical environment in mind and this varies with time from the Baroque to the Romantic to the Modern musical period. Conductors vary their interpretation according to the hall they are in. Well-traveled listeners and music critics have favorite halls. The lecture then presents a list of 58 halls rank ordered according to their acoustical quality based on interviews of music critics and conductors. Modern acoustical measurements made in these halls are compared with their rankings. Music recordings will be presented that demonstrate how halls sound that have different measured acoustical parameters. Photographs of a number of recently built halls are shown as examples of how these known acoustical factors have been incorporated into architectural design.

Session 2aAAa

Architectural Acoustics: Methods to Quantify Opera House Acoustics I: Basic Studies

Roberto Pompoli, Chair
Dipartimento di Ingegneria, Universita di Ferrara, Via G Saragat 1, 44100 Ferrara, Italy

Chair's Introduction—7:30

Invited Papers

7:40

2aAAa1. Theory of acoustic design of opera house and a design proposal. Yoichi Ando (1-10-27 Yamanokami, Kumamoto 862-0915, Japan)

First of all, the theory of subjective preference for sound fields based on the model of auditory-brain system is briefly mentioned. It consists of the temporal factors and spatial factors associated with the left and right cerebral hemispheres, respectively. The temporal criteria are the initial time delay gap between the direct sound and the first Reflection (Dt1) and the subsequent reverberation time (Tsub). These preferred conditions are related to the minimum value of effective duration of the running autocorrelation function of source signals (te)min. The spatial criteria are binaural listening level (LL) and the IACC, which may be extracted from the interaural crosscorrelation function. In the opera house, there are two different kind of sound sources, i.e., the vocal source of relatively short values of (te)min in the stage and the orchestra music of long values of (te)min in the pit. For these sources, a proposal is made here.

8:00

2aAAa2. Magnetoencephalographic responses in relation to temporal and spatial factors of sound fields. Yoshiharu Soeta, Seiji Nakagawa, Mitsuo Tonoike (Life Electron. Lab., Natl. Inst. of Adv. Industrial Sci. and Technol. (AIST), Midorigaoka, Ikeda, Osaka 563-8577, Japan), Takuya Hotehama, and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe 657-8501, Japan)

To establish the guidelines based on brain functions for designing sound fields such as a concert hall and an opera house, the activities of the human brain to the temporal and spatial factors of the sound field have been investigated using magnetoencephalography (MEG). MEG is a noninvasive technique for investigating neuronal activity in human brain. First of all, the auditory evoked responses in change of the magnitude of the interaural cross-correlation (IACC) were analyzed. IACC is one of the spatial factors, which has great influence on the degree of subjective preference and diffuseness for sound fields. The results indicated that the peak amplitude of N1m, which was found over the left and right temporal lobes around 100 ms after the stimulus onset, decreased with increasing the IACC. Second, the responses corresponding to subjective preference for one of the typical temporal factors, i.e., the initial delay gap between a direct sound and the first reflection, were investigated. The results showed that the effective duration of the autocorrelation function of MEG between 8 and 13 Hz became longer during presentations of a preferred stimulus. These results indicate that the brain may be relaxed, and repeat a similar temporal rhythm under preferred sound fields.

8:20

2aAAa3. Acoustical measurements of sound fields between the stage and the orchestra pit inside an historical opera house. Shin-ichi Sato, Nicola Prodi (Eng. Dept., Univ. of Ferrara, via Saragat 1, 44100 Ferrara, Italy), and Hiroyuki Sakai (Kobe Univ., Rokkodai, Nada, Kobe 657-8501, Japan)

To clarify the relationship of the sound fields between the stage and the orchestra pit, we conducted acoustical measurements in a typical historical opera house, the Teatro Comunale of Ferrara, Italy. Orthogonal factors based on the theory of subjective preference and other related factors were analyzed. First, the sound fields for a singer on the stage in relation to the musicians in the pit were analyzed. And then, the sound fields for performers in the pit in relation to the singers on the stage were considered. Because physical factors vary depending on the location of the sound source, performers can move on the stage or in the pit to find the preferred sound field.

Contributed Papers

8:40

2aAAa4. A scale value for the balance between stage and pit and inside an historical opera house. Nicola Prodi and Sylvia Velecka (Dipartimento di Ingegneria, Universit di Ferrara, via Saragat1, Ferrara, Italy)

Despite its recognized importance, the balance between the singer and the orchestra inside an opera house has received minor attention in the past. In fact, after the fundamental work of Meyer [J. Meyer, ''Some problems of opera house acoustics,'' Proceedings of 12th I.C.A., Vancouver, 1986, pp. 13–18], who explained why the solo singing voice can compete with the orchestra, only partial results were reported on this perceived attribute. In this work a reference scale to assess the balance inside an historical opera house is achieved by means of listening tests inside a controlled room. Two scaling experiments were performed based on the acoustical data measured inside an historical opera house, the Teatro Comunale di Ferrara, Italy. By doing so all of the relevant acoustical characteristics of a typical Italian-style opera house could be exactly reproduced. Acceptable values do not differ much in the stalls and in the boxes and are within 2 dB(A) to +2.3 dB(A). The transfer of the findings to other types of opera houses is discussed, too.

8:55

2aAAa5. Analysis of masking effects on speech intelligibility with respect to moving sound stimulus. Chiung Yao Chen (168, Gifeng E. Rd., Wufeng, Taichung County 413, Taiwan, R.O.C., chychen@mail.cyut.edu.tw)

The purpose of this study is to compare the disturbed degree of speech by an immovable noise source and an apparent moving one (AMN). In the study of the sound localization, we found that source-directional sensitivity (SDS) well associates with the magnitude of interaural cross correlation (IACC). Ando *et al.* [Y. Ando, S. H. Kang, and H. Nagamatsu, J. Acoust. Soc. Jpn. (E) **8**, 183–190 (1987)] reported that potential correlation between left and right inferior colliculus at auditory path in the brain is in harmony with the correlation function of amplitude input into two ear-canal entrances. We assume that the degree of disturbance under the apparent moving noisy source is probably different from that being installed in front of us within a constant distance in a free field (no reflection). Then, we found there is a different influence on speech intelligibility between a moving and a fixed source generated by 1/3-octave narrowband noise with the center frequency 2 kHz. However, the reasons for the moving speed and the masking effects on speech intelligibility were uncertain.

9:10

2aAAa6. Multi-dimensional analysis of subjective acoustical ratings and acoustical measures in existing concert halls. Toshiyuki Okano (Takenaka Res. and Development Inst., 1-5-1, Ohtsuka, Inzai, Chiba 270-1395, Japan, okano.toshiyuki@takenaka.co.jp)

Correlations between subjective acoustical ratings and hall-averaged values of acoustical measures are studied among existing worldwide major concert halls. It was shown that the classified acoustical ratings by Beranek [*Concert and Opera Halls, How They Sound* (ASA, 1996)] are discriminated correctly by combining binaural quality index (BQI) with some other acoustical measures. BQI is determined by the arithmetic average of inter-aural cross correlation coefficient in three octave bands of 500, 1000, and 2000 Hz, subtracted from unity, calculated from the early 80-ms part of binaural impulse response. Considering that the upper limit value of BQI not to cause disturbing image shift is approximately 0.85 at individual seat [Okano, J. Acoust. Soc. Am. 2219–2230 (2000)], the values of 0.6 or higher in hall averaged value of BQI, 0.85 or smaller in individual seat value of BQI, and approximately 5 dB or higher in strength factor at middle frequencies are proposed as design objectives to attain a high acoustical quality. It should be provided that other acoustical measures are also optimized. These target values will be very effective in studying room shape of halls, using scale models or computer models.

10:20

2aAAa7. Microperforated stretched ceilings basics and applications. Christian Nocke and Catja Hilge (Akustikbuero Oldenburg, Alte Raad 20a, D-26127 Oldenburg, Germany, nocke@akustikbuero.info)

In 2001 a new product based on the theory of microperforated sound absorbers has been introduced. The thin foil of a stretched ceiling has been microperforated. Stretched-membrane ceilings have been a popular product with regard to modern architecture worldwide for more than 25 years. However, so far only optical and other aspects of the product were generally of interest. A brief introduction of the underlying theory of microperforated sound absorbers according to D.-Y. Maa will be given. Four geometrical parameters (diameter of perforation, distance between holes, thickness of panel/foil, and backing air volume) fully determine the sound-absorption coefficient of these sound absorbers. Results of independent qualification measurements of various assemblies will be shown. The agreement between theoretical predictions and measurements from the reverberation chamber is very high. Finally, different applications in a church, a sports hall, and other rooms will be presented. Again, the comparison between theoretical design and final result show a very high degree of agreement.

9:40

2aAAa8. Lateral efficiency in small auditoriums. Henrik Moller and Timo Peltonen (Akukon Oy Consulting Engineers, Kornetintie 4 A, FIN-00380 Helsinki, Finland, henrik.muller@akukon.fi)

The paper will examine the measured lateral efficiency parameters measured in small halls, that is, halls with less than 750 seats. Comparisons between measured LF and IACC parameters will be presented as well as a study of the correlation between the lateral efficiency parameters and the width and other geometrical attributes of the halls. The measurements were done as part of a larger survey of the Finnish halls that were built between 1980 and 2000.

TUESDAY MORNING, 25 MAY 2004 ROYAL BALLROOM A, 10:15 TO 11:55 A.M.

Session 2aAAb

Architectural Acoustics: Methods to Quantify Opera House Acoustics II: Singer Acoustics

Yoichi Ando, Chair
1-10-27 Yamanokami, Kumamoto 860-0915, Japan

Chair's Introduction—10:15

Invited Papers

10:20

2aAAb1. Singer's preferred acoustic condition in performance in an opera house and self-perception of the singer's voice. Dennis Noson (BRC Acoust., 1741 1st Ave. S., Ste. 401, Seattle, WA 98134, noson@alum.mit.edu), Kosuke Kato,[a] and Yoichi Ando[b] (Kobe Univ., Kobe, Japan)

Solo singers have been shown to over estimate the relative sound pressure level of a delayed, external reproduction of their own voice, singing single syllables, which, in turn, appears to influence the preferred delay of simulated stage reflections [Noson, Ph.D. thesis, Kobe University, 2003]. Bone conduction is thought to be one factor separating singer versus instrumental performer judgments of stage acoustics. Using a parameter derived from the vocal signal autocorrelation function (ACF envelope), the changes in singer

preference for delayed reflections is primarily explained by the ACF parameter, rather than internal bone conduction. An auditory model of a singer's preferred reflection delay is proposed, combining the effects of acoustical environment (reflection amplitude), bone conduction, and performer vocal overestimate, which may be applied to the acoustic design of reflecting elements in both upstage and forestage environments of opera stages. For example, soloists who characteristically underestimate external voice levels (or overestimate their own voice) should be provided shorter distances to reflective panels—irrespective of their singing style. Adjustable elements can be deployed to adapt opera houses intended for bel canto style performances to other styles. Additional examples will also be discussed. [a] Now at Kumamoto Univ., Kumamoto, Japan. [b] Now at: 1-10-27 Yamano Kami, Kumamoto, Japan.

10:40

2aAAb2. Blending vocal music with a given sound field due to the characteristics of the running autocorrelation function of singing voices. Kosuke Kato (Grad. School of Sci. and Technol., Kumamoto Univ., 2-39-1 Kurokami, Kumamoto, Japan, 031d9203@gsst.stud.kumamoto-u.ac.jp), Kenji Fujii (Yoshimasa Electron., Inc., Shibuya-ku, Tokyo, Japan), Keiji Kawai (Kumamoto Univ., Kumamoto, Japan), Yoichi Ando[a] (Kobe Univ., Kobe, Japan), and Takashi Yano (Kumamoto Univ., Kumamoto, Japan)

This is a study to meet music and the opera house acoustics. It is said that singers adjust their interpretation style according to the acoustical condition of the sound field in a room. However, this mechanism of blending of musical performance with the sound field is unknown. In order to obtain a method of performance blending of opera house acoustics, we attempted to develop evaluation criteria for a singing voice in terms of the minimum value of the effective duration of the running autocorrelation function (r-ACF), (te)min, of sound source signals. This temporal factor has shown to have close correlation with the subjective response of both listeners and performers to sound fields [Y. Ando, *Architectural Acoustics* (AIP Press/Springer-Verlag, New York, 1998)]. As example for the control of (te)min due to performing style, effects of singing style, kind of vowel, relative pitch, vibrato extent, and intonation on the values of (te)min are demonstrated. In addition, the fine structure of the r-ACF is discussed with regard to the identification of vowels of singing voice. [a] Now at 1-10-27 Yamanokami, Kumamoto, Japan.

11:00

2aAAb3. Acoustical qualification of Teatro Nuovo in Spoleto before refurbishing works. Alessandro Cocchi, Marco Cesare Consumi, and Ryota Shimokura (DIENCA, Eng. Faculty, Viale Risorgimento 2, 40136 Bologna, Italy)

To qualify the acoustical quality of an opera house two different approaches are now available: one is based on responses of qualified listeners (subjective judgments) compared with objective values of selected parameters, the other on comparison tests conducted in suited rooms and on a model of the auditory brain system (preference). In the occasion of the refurbishment of an opera house known for the Two Worlds Festival edited yearly by the Italian Composer G. C. Menotti, a large number of measurements were taken with different techniques, so it is possible to compare the different methods and also the results with some geometrical criterion, based on the most simple rules of musical harmony, now neglected as our attention is attracted to computer simulations, computer aided measurement techniques and similar modern methods. From this work some link between well known acoustical parameters (not known at the time when architects sketched the shape of ancient opera houses) and geometrical criteria (well known at the time when ancient opera houses were built) will be shown.

11:20

2aAAb4. Proposal of a new plane shape of an opera house—optimized by genetic algorithms. Takuya Hotehama, Yoichi Ando, Akinori Tani, and Hiroshi Kawamura (Grad. School of Sci. and Technol., Kobe Univ.. Rokkodai, Nada, Kobe 657-8501, Japan, 002d885n@y02.kobe-u.ac.jp)

The horseshoe-shaped theater has been the main shape from historical circumstances. However, from acoustical points of view, the rationality of the peculiar plane shape is not yet verified more than historical refinement. In this study, in order to make the theater shape more acoustically excellent, optimization for temporal and spatial factors in the theory of the subjective preference was made using genetic algorithms (GAs) by operating the positions of side walls. Results reconfirm that the plane shape of the optimized theater is a leaf shape, which has been verified to be acoustically rational in a concert hall. And, further possible shapes are also offered.

Contributed Paper

11:40

2aAAb5. The control of balance between singer on the stage and orchestra in the pit by means of virtual opera house models. Linda Parati, Roberto Pompoli, and Nicola Prodi (Dept. of Eng., Univ. of Ferrara, Via Saragat 1, 44100 Ferrara, Italy, lparati@ing.unife.it)

In this study some of the architectural elements which can influence the balance between the singer on the stage and the orchestra in the pit, in respect to the design constraints, have been investigated. By means of room simulations, modifications in the type of materials, dimensions, and shapes of the pit, and modifications of the proscenium dimension and stage slope were applied. Some of the architectural elements controlling the balance have been identified and optimized. Simulations were performed with omnidirectional sources in the pit, despite the directivity of the instruments, and directional source on the stage. The effects of the applied changes on the balance between singer, conductor, and orchestra have been analyzed too.

Session 2aAAc

Architectural Acoustics: Theaters for Drama Performance—Another Two Decades (1984–2004) (Poster Session)

Robin Glosemeyer, Chair
Jaffe Holden Acoustics, Inc., 501 Santa Monica Boulevard, Santa Monica, California 90401

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

2aAAc1. Albert Ivar Goodman Theatre, The Goodman Theatre. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc2. Bingham Theatre, Actors Theatre of Louisville. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc3. Center Theater, California Center for the Arts. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc4. Chicago Shakespeare Theater. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc5. Lookingglass Theatre. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc6. Steppenwolf Theatre. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc7. Owen Bruner Goodman Theatre, The Goodman Theatre. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc8. The O'Reilly Theater, Pittsburgh Public Theater. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc9. The Allen Theatre, ACT Theatre. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc10. The Falls Theatre, ACT Theatre. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc11. Fichandler Stage, Arena Stage. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc12. The Cradle, Arena Stage. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc13. Sidney Harman Hall, Harman Center for the Arts. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc14. Second Theatre, Children's Theatre Company. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc15. Thrust Stage, Guthrie on the River. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc16. Proscenium Stage, Guthrie on the River. Richard H. Talaske (The Talaske Group, Inc., 105 N. Oak Park Ave., Oak Park, IL 60301, rick@talaske.com)

2aAAc17. Rodgers Victory Theatre. Steven Thorburn (P.O. Box 20399, Castro Valley, CA 94546, sjt@ta-inc.com)

2aAAc18. Drama Theater of the New National Theater (NNT), Tokyo, Japan. Takayuki Hidaka (Takenaka R&D, 1-5-1, Otsuka, Inzai, Chiba 270-1395, Japan) and Leo L. Beranek (Cambridge, MA 02138-5755)

2aAAc19. Roger S. Berlind Theatre, McCarter Theatre Center, Princeton, NJ. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2aAAc20. New Amsterdam Theater. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc21. Purnell Center for the Arts, Carnegie Mellon University. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc22. New Second Stage Theatre. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc23. Biltmore Theater. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc24. Sarofim Hall, The Hobby Center for the Performing Arts. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc25. Yale Drama Theater. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc26. Majestic Hall, Brooklyn Academy of Music. Robin Glosemeyer (Jaffe Holden Acoustics, Inc., 501 Santa Monica Blvd., Ste. 606, Santa Monica, CA 90401, rglosemeyer@jhacoustics.com)

2aAAc27. Experimental theater of the New National Theater Tokyo, Japan. Takayuki Hidaka (Takenaka R&D, 1-5-1, Otsuka Inzai, Chiba 270-1395, Japan, hidaka.takayuki@takenaka.co.jp) and Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755)

2aAAc28. Kokos, The Theater Academy of Finland. Henrik Moller and Tarja Lahti (Akukon Oy Consulting Engineers, Kornetintie 4 A, FIN-00380 Helsinki, Finland, henrick.moller@akukon.fi)

2aAAc29. Oulu City Theater. Henrik Moller and Anssi Ruusuvuori (Akukon Oy Consulting Engineers, Kornetintie 4 A, FIN-00380 Helsinki, Finland, henrik.moller@akukon.fi)

2aAAc30. Neal-Marshall Education Center, Indiana University. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2aAAc31. Cultural Life Center, Roberts Wesleyan College. R. Kring Herbert (Ostergaard Acoustical Assoc., 200 Executive Dr., West Orange, NJ 07052, kherbert@acousticalconsultant.com)

2aAAc32. Acoustical design of the Bayside Performing Arts Center. Jason R. Duty and David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc33. Acoustical design of the Randall Jr. Museum Theater. David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc34. Acoustical design of the Mendocino Community College Fine Arts Theatre. David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc35. Vacaville Community Theater acoustics. Kenneth W. Graven and David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc36. Sound isolation design for REDCAT (Roy and Edna Disney/CalArts Theater). Cristina L. Miyar, Thomas A. Schindler, and David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc37. Newburyport Firehouse Theater. William J. Cavanaugh and Timothy J. Foulkes (Cavanaugh Tocci Assoc., Inc., cstorch@cavtocci.com)

2aAAc38. Capitol Center Theater. William J. Cavanaugh and K. Anthony Hoover (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA, cstorch@cavtocci.com)

2aAAc39. Westminster School Fine Arts Center. Timothy J. Foulkes (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA, cstorch@cavtocci.com)

2aAAc40. Murat Theater and Temple. William J. Cavanaugh (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA, cstorch@cavtocci.com)

2aAAc41. Pinkerton Academy Theater. Lincoln B. Berry (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA, cstorch@cavtocci.com)

2aAAc42. Rhode Island College Music Building. Lincoln B. Berry (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA, cstorch@cavtocci.com)

2aAAc43. Neptune Theater. Peter Terroux (Atlantic Acoust. Assoc.), Gregory C. Tocci, and Christopher A. Storch (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA, cstorch@cavtocci.com)

2aAAc44. Ship's Company Theater. Peter Terroux (Atlantic Acoust. Assoc., P.O. Box 2520, DEPS, Dartmouth, NS B2W 4A5, Canada)

2aAAc45. Restoration of the Geary Theatre, San Francisco. Ewart A. Wetherill (Ewart A. Wetherill AIA, 28 Cove Rd., Alameda, CA 94502, redwetherill@sbcglobal.net)

2aAAc46. San Jose Repertory Theater acoustics. Philip N. Sanders and David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc47. Acoustical design for the Fairfield Center for Creative Arts. David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2aAAc48. Blodgett Hall Auditorium, Vassar College. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2aAAc49. Acoustical design of the ACM Theater ATM. Hideo Nakamura and Keiji Oguchi (Nagata Acoust., Inc., Hongo-Segawa Bldg. 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan, nakamura@nagata.co.jp)

2aAAc50. The Theatre, The Carlsen Center, Johnson County Community College, Overland Park, Kansas. Robert C. Coffeen (School of Architecture and Urban Design, The Univ. of Kansas, Lawrence, KS 66045)

2aAAc51. Saitama Arts Theater (Main Theater). Shinichi Sawara, Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, sawara@beat.yamaha.co.jp), and Kiyoteru Ishii (Univ. of Tokyo, Tokyo, Japan)

2aAAc52. Hakata-za. Takayuki Watanabe, Takashi Yamakawa, and Yasushi Shimizu (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, watanabe@beat.yamaha.co.jp)

2aAAc53. Nagakute Cultural Center Mori-no Hall. Masato Hata, Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, hata@beat.yamaha.co.jp), Yuji Sonoda, and Hideki Tachibana (Inst. of Industrial Sci., Univ. of Tokyo)

2aAAc54. Arkas SASEBO. Masahiro Ikeda, Shinji Kishinaga, and Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, ikeda@beat.yamaha.co.jp)

2aAAc55. Kani Public Arts Center. Masahiro Ikeda, Shinji Kishinaga, Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, ikeda@beat.yamaha.co.jp), and Masahito Yasuoka (Tokyo Univ. of Sci., Tokyo, Japan)

2aAAc56. Noto Engekido. Takashi Yamakawa, Masato Hata, and Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan)

2aAAc57. National Theater Okinawa. Takashi Yamakawa, Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan), Yukiya Tokuriki (Takamatsu Architects and Assoc. Co. Ltd., Osaka, Japan), Kiyoshi Masuda (Taisei Corp., Yokohama, Japan), and Kiyoteru Ishii (Univ. of Tokyo, Tokyo, Japan)

2aAAc58. Roschel Performing Arts Center, Franklin & Marshall College. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2aAAc59. Biwako Hall, Center for the Performing Arts, Shiga. Shinji Kishinaga, Rento Tanase, and Fukushi Kawakami (Yamaha Corp., Advance System Development Ctr., Hamamatsu, Japan, kishinaga@beat.yamaha.co.jp)

2aAAc60. Nohgaku-do, Niigata City Performing Arts Center. Yasushi Shimizu, Yoshikazu Honji, and Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, shimizu@beat.yamaha.co.jp)

2aAAc61. Theater, Niigata City Performing Arts Center. Yasushi Shimizu, Yoshikazu Honji, and Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, shimizu@beat.yamaha.co.jp)

2aAAc62. Civic Theater in Shonandai Culture Center. Yasushi Shimizu, Shinjiro Yamashita, and Fukushi Kawakami (Yamaha Corp., Adv. System Development Ctr., Hamamatsu, Japan, shimizu@beat.yamaha.co.jp)

2aAAc63. Act City Hamamatsu Main Hall. Takayuki Watanabe, Shinji Kishinaga, and Fukushi Kawakami (Yamaha Corp. Adv. System Development Ctr., Hamamatsu, Japan, watanabe@beat.yamaha.co.jp)

2aAAc64. Cultural Arts Center, Baton Rouge, LA. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2aAAc65. Hasty Pudding Theatre, Harvard University. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2aAAc66. The Swisher Theater. Gary Siebein, Hyun Paek, Stephen Skorski, Robert Lilkendey (Siebein Assoc., Inc., 625 NW 60th St., Ste. C, Gainesville, FL 32607, gsiebein@siebeinacoustic.com), and Mac McGowan (McGowan Sound Design, Jacksonville, FL)

2aAAc67. Gulfport Community Theater. Gary Siebein, Hyun Paek, Stephen Skorski (Siebein Assoc., Inc., 625 NW 60th St., Ste. C, Gainesville, FL, gsiebein@siebeinacoustic.com), and Michael Ermann (Virginia Polytechnic and State Univ.)

2aAAc68. The Players' Theater. Gary Siebein, Stephen Skorski, Hyun Paek (Siebein Assoc., Inc., 625 NW 60th St., Ste. C, Gainesville, FL 32607, gsiebein@siebeinacoustic.com), and Martin Gold (School of Architecture, Univ. of Florida)

2aAAc69. Horace Mann. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, New York 10801, info@acousticdimensions.com)

2aAAc70. Illinois State University Theater. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, New York 10801, info@acousticdimensions.com)

2aAAc71. Berrie Center—Ramapo College. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

TUESDAY MORNING, 25 MAY 2004 IMPERIAL BALLROOM B, 9:00 A.M. TO 12:00 NOON

Session 2aAB

Animal Bioacoustics: Animal Bioacoustics Posters

Whitlow W. L. Au, Chair

Hawaii Institute of Marine Biology, P.O. Box 1106, Kailua, Honolulu 96734

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.

2aAB1. Hearing in the North Atlantic right whale: Anatomical predictions. Susan E. Parks,[a] Darlene R. Ketten, Jennifer Trehey O'Malley, and Julie Arruda (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, sparks@whoi.edu)

Understanding the hearing abilities of large whales is important to determine the impacts of anthropogenic sources of sound. Right whales are not amenable to traditional physiological techniques to test hearing. Previous research on the hearing of marine mammals has shown that functional morphometric models are reliable estimators of hearing sensitivity in marine species. Morphometric analyses of 18 inner ears from 13 stranded right whales were used in the development of a preliminary model of the frequency range of hearing. All ears were scanned with computerized tomography (CT). Four ears showing the best preservation were processed into slides for measurements of the basilar membrane. Calculated basilar-membrane length averaged 55.7 mm (range 50.5–61.7 mm). The ganglion cell density/mm averaged 1842 ganglion cells/mm. Membrane length and ganglion cell density were used to predict a total ganglion cell count of approximately 102 500 ganglion cells for right whales. The thickness/width measurements of the basilar membrane from slides resulted in an estimated frequency range of approximately 10 Hz–22 kHz based on established marine mammal models. Additional measurements from more specimens will be necessary to develop a more robust model of the right whale hearing range. [a]Currently at Cornell University Bioacoustics Research Program.

2aAB2. Song variation and environmental auditory masking in the grasshopper sparrow. Bernard Lohr, Robert J. Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742, blohr@psyc.umd.edu), and Douglas E. Gill (Univ. of Maryland, College Park, MD 20742)

Some grassland bird species, in particular grasshopper sparrows (*Ammodramus savannarum*), sing songs with especially high mean frequencies (7.0–8.0 kHz). Acoustic interference is one potential explanation for the evolution of high frequency vocalizations, particularly in open habitats. We tested predictions from a model of effective auditory communication distances to understand the potential effects of vocal production and environmental auditory masking on vocal behavior and territoriality. Variation in the spectral structure of songs and the size and shape of territories was measured for grasshopper sparrows in typical grassland habitats. Median territory areas were 1629 m^2 at a site in the center of the species range in Nebraska, and 1466 m^2 at our study site in Maryland,

with average territory diameters measuring 20.2 m. Species densities and sound pressure levels also were determined for stridulating insects and other noise sources in the habitat. Based on current models of effective communication distances, known noise levels, and information on hearing abilities, our results suggest that auditory sensitivity and environmental noise could be factors influencing the mean frequency and spatial dynamics of territorial behavior in grassland birds. [Work supported by NIH and the CRFRC.]

2aAB3. Reduced temporal integration in Belgian Waterslager canaries. Amanda M. Lauer and Robert J. Dooling (Psych. Dept. and Ctr. for Comparative and Evolutionary Biol. of Hearing, Univ. of Maryland, College Park, MD 20743)

Belgian Waterslager canaries (BWS) are bred for their low-frequency song and have been shown to have hair cell abnormalities in the inner ear that result in elevated thresholds at higher frequencies. Previous results show that resolution of temporal fine structure, or the ability to resolve rapidly occurring changes in complex sounds, is enhanced in BWS canaries. In a continuing effort to assess the effects of the BWS inner ear pathology on hearing, we here investigate the ability to integrate acoustic information over longer periods of time. Absolute thresholds for 1.0-, 2.0-, and 4.0-kHz pure tones were measured in BWS and normal-hearing non-BWS canary strains for durations ranging from 5 to 480 ms using operant conditioning methods. NonBWS canaries showed a decrease in threshold of approximately 10–15 dB with increasing tone duration for all frequencies. In contrast, BWS showed almost no change in threshold across the range of durations tested for all frequencies. The reduced temporal integration in BWS canaries with hair cell abnormalities parallels similar findings in humans with cochlear damage. [Work supported by NIH DC01372 to RJD and DC05450 to AML.]

2aAB4. Changing the average frequency of contact calls is associated with changes in other acoustic parameters in the budgerigar (*Melopsittacus undulatus*). Michael Osmanski and Robert Dooling (Dept. of Psych., Univ. of Maryland, College Park, MD 20742)

The most-often produced vocalization of the budgerigar, a small parrot native to Australia, is the short (100–150 ms) frequency-modulated contact call. These calls play a role in maintaining flock dynamics and are believed to act as vocal signatures in these birds. Previous findings in our lab have shown that budgerigars can control the intensity of their vocal behavior and exhibit a robust Lombard effect (Manabe *et al.*, 1998). Recently, we have shown that there is a high degree of stereotypy in contact calls across a number of acoustic parameters (Osmanski and Dooling, 2004). Questions arise concerning the limits of plasticity in these calls and the relation or interdependence among the various parameters. As a first approach to answering these questions, four budgerigars were trained using operant conditioning methods to change the average peak frequency of their contact calls (both upward and downward in frequency) to obtain access to a food reward. Results show that these birds can both increase and decrease the average frequency of their contact calls. Such changes are associated with modifications in a number of other acoustic parameters, suggesting constraints on vocal plasticity. [Work supported by NIH DC-00198 to RJD and NIDCD Training Grant DC-00046.]

2aAB5. An audiometric comparison of primate audiograms. Mark N. Coleman (Dept. of Anthropology, Stony Brook Univ., Stony Brook, NY 11794)

Audiogram data for 18 species of primates were collected from the literature and analyzed by measuring 13 audiometric variables: frequency and threshold of the primary peak, frequency and threshold of the secondary peak, frequency and threshold of the notch between peaks, low-frequency cutoff, high-frequency cutoff, total area of the audible field, low area, middle area, high area, and total audible range in octaves. All areal measurements were made using IGOR PRO 4.04 wave measurement software. Platyrrhines were found to have significantly better low-frequency sensitivity than like-sized lorisoids with an average of 15-dB difference between the means for the two groups. This difference remains significant even when interindividual variation is considered. Callithrix jacchus and Erythrocebus patas have unusual hearing patterns for primates of their size with marmosets showing a reduction in high-frequency sensitivity, while patas monkeys show a reduction in low-frequency sensitivity. It was also noted that chimps have a notch in sensitivity that falls within the range of greatest sensitivity for humans. These findings are discussed in relation to the morphological adaptations that appear to influence these hearing patterns and the evolutionary significance of such patterns for group communication and predator–prey interactions.

2aAB6. Influence of low-frequency vibration on changes of biochemical parameters of living rats. Cezary Kasprzak, Zbigniew Damijan, and Ryszard Panuszka (Univ. of Sci. and Technol., Lab. of Structural Acoust. and Biomed. Eng., Al. Mickiewicza 30, 30-059 Krakow, Poland, cekasp@agh.edu.pl)

The aim of the research was to investigate how some selected biochemical parameters of living rats depend on exposure of low-frequency vibrations. Experiments were run on 30 Wistar rats randomly segregated into three groups: (I) 20 days old (before puberty), (II) 70th day after; (III) control group. The exposure was repeated seven times, for 3 h, at the same time of day. Vibrations applied during the first tests of the experiment had acceleration 1.22 m/s^2 and frequency 20 Hz. At the 135th day the rats' bones were a subject of morphometric/biochemical examination. The results of biochemical tests proved decrease in LDL and HDL cholesterol levels for exposed rats as well as the Ca contents in blood plasma. There was evident increasing of Ca in blood plasma in exposed rats for frequency of exposition.

Session 2aAO

Acoustical Oceanography: Ocean Acoustics of Earthquakes

Ralph A. Stephen, Cochair
Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543-1542

Robert I. Odom, Cochair
Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105-6698

Chair's Introduction—8:00

Invited Papers

8:05

2aAO1. Characteristics of *T* phases from earthquake sources, or how earthquakes "talk" into the ocean. Emile A. Okal (Dept. of Geol. Sci., Northwestern Univ., Evanston, IL 60208, emile@earth.nwu.edu)

We review the mechanisms by which earthquake sources within the solid Earth generate acoustic energy into the SOFAR channel. We examine individual cases involving both the relatively inefficient mechanism of ''downslope conversion,'' leading to *T* phases with low amplitudes and long durations, and efficient conversion at steep interfaces allowing direct penetration of the SOFAR, larger amplitudes, and shorter durations. We derive several ways to quantify *T* waves generated by an earthquake source. To avoid the saturation of their amplitude as source size grows, we use the concept of a *T*-phase energy flux (TPEF), which mirrors the radiated energy measured on conventional seismic body waves. TPEF is scaled to the seismic moment of the earthquake to define a parameter, gamma, characterizing the earthquake's efficiency for *T*-phase generation. In another approach, we compare the amplitude of *T*-phase envelopes to the duration of their signal. The resulting discriminant has been proposed to identify earthquakes from explosions. We show that it varies for several categories of nonstandard earthquakes, notably those occurring in volcanic islands and events with slow rupture giving rise to large tsunamis. Finally, we discuss a few examples of *T* waves from underwater landslides.

8:25

2aAO2. Wave-theory modeling of oceanic *T*-phase coupling at continental margins and seamounts. Henrik Schmidt, Arthur B. Baggeroer (Dept. of Ocean Eng., MIT, Cambridge, MA 02139), and Brian J. Sperry (Sci. Application Intl. Corp., McLean, VA 22102)

The role of seismo-acoustic seabed scattering as a mechanism for coupling of seismic energy into oceanic teleseismic waves or *T*-phases is investigated using a new versatile modeling capability for seismo-acoustic propagation in laterally inhomogeneous or range-dependent ocean waveguides. The Virtual Source Approach (VISA) uses a local Rayleigh–Kirchhoff approximation to handle the transmission and reflection of plane waves at the vertical interfaces separating horizontally ocean stratified sectors. Combined with the wavenumber integration approach which inherently computes the plane-wave decomposition of the seismo-acoustic field in stratified fluid-elastic waveguides, this approach provides a robust approximation to the seismo-acoustic coupling phenomena in shallow and deep ocean waveguides. The VISA approach has been implemented in the OASES seismo-acoustic modeling framework and used to investigate the role of seismo-acoustic conversion and scattering by seabed topography and roughness in generating oceanic *T*-phases at continental margins and seamounts. It is demonstrated that the excitation of the oceanic *T*-phases can be explained by the coupling of crustal shear body-waves into seismic interface waves, or seabed Scholte waves, which then subsequently scatter into the waterborne modal spectrum. This wavenumber conversion mechanism implies that the excitation of the *T*-phases will be significantly stronger by earthquakes producing crustal SV-waves than those producing predominantly P-waves. This in turn suggests that earthquakes associated with dip-slip failure modes excite significantly stronger *T*-phases than buried explosive sources.

8:45

2aAO3. Computation of *T*-phase seismograms using stairsteps to model seafloor bathymetry changes. Catherine de Groot-Hedlin (Scripps Inst. of Oceanogr., UCSD, 9500 Gilman Dr., La Jolla, CA 92093, chedlin@ucsd.edu)

Stair-step boundaries are employed to represent range-dependent seafloor bathymetry in many types of numerical modeling methods. However, if an overly coarse discretization is used to simulate smoothly varying bathymetry, the acoustic solution is degraded in a way that simulates scattering. Geometrical optics approximations can be used to derive discretization criteria for simulating a smoothly sloping interface for the case of a source embedded in either an acoustic or an elastic seafloor, and applied to modeling *T* phases. A finite difference time-domain modeling approach is used to synthesize *T* phases for both smoothly sloping and rough seafloor boundaries. It is shown that scattering at a rough seafloor boundary yields ocean-borne acoustic phases with velocities near those of observed *T* phase, while smooth seafloor models yield *T* phases with slower horizontal velocities. The long duration of the computed *T* phases for both the rough acoustic and elastic models is consistent with energy being scattered into the sound channel both as it transits the ocean/crust boundary, as well as at several subsequent seafloor reflections. However, comparison between the elastic and acoustic modeling solutions indicates that the *T*-phase wavetrain duration decreases with decreasing impedance contrast between the ocean and seafloor.

9:05

2aAO4. The use of regional hydrophone datasets for statistical studies of mid-ocean ridge earthquake populations. DelWayne R. Bohnenstiehl (Lamont-Doherty Earth Observatory of Columbia Univ., P.O. Box 1000, Palisades, NY 10964)

Seismically generated tertiary (T-) waves can be used to detect and locate the shallow hypocenter earthquakes associated with mid-ocean ridge spreading centers and oceanic transforms. During the last several years, regional seismic catalogs, which contain estimates of T-wave source regions, origin times and acoustic source levels, have been generated from continuous hydroacoustic monitoring along the northern Mid-Atlantic Ridge and equatorial East Pacific Rise. As the mechanics of T-wave generation remain poorly understood, critics of these datasets have questioned the correlation between T-wave source regions and seismic epicenters, as well as the usefulness of acoustic source level as a proxy for earthquake size. Recent analyses, however, suggest that these data provide a reasonably complete record of seismicity within the mid-ocean ridge environment, yielding significant improvements in location accuracy and catalog completeness level, relative to global seismic catalogs. Most importantly, it appears that many fundamental and well-established properties of seismic distributions can be studied using T-wave catalogs, including the modified Omori decay law of aftershocks, a power-law size frequency distribution (Gutenberg–Richter law), a fractal time-clustering behavior ($1/f$ noise) and fractal spatial distribution of epicenters. These observations show that T-wave derived earthquake catalogs can be used for quantitative seismo-tectonic studies in the ridge setting.

9:25

2aAO5. Hydroacoustic observations of sub-Indian ocean earthquakes using arrays of the International Monitoring System. Jay J. Pulli and Zachary M. Upton (BBN Technologies, 1300 N. 17th St., Ste. 400, Arlington, VA 22209, jpulli@bbn.com)

Until recently, there have been few public hydroacoustic recordings of earthquakes beneath the Indian Ocean. Now, the International Monitoring System, which includes three Indian Ocean station sites, is providing excellent data for study. These stations consist of three-element horizontal hydrophone arrays floated to the middle of the sound channel axis, and can record low-level signals from distant sources. System bandwidth is approximately 2–100 Hz. We have accumulated a waveform database of 150 sub-Indian Ocean earthquakes and surrounding continental events recorded by this system. These data allow us to research T-wave generation and propagation. The recordings include seismic body-wave conversions beneath the arrays, T waves, and long-range reflections. Recordings of main shock/aftershock series have allowed us to compute scaling relationships for T-wave amplitudes versus magnitude and examine T-wave excitation versus source mechanism. When combined with distance corrections, we can scale T waves recorded at a given station to a fixed magnitude and distance. Plotted versus azimuth around the station, these scaled T-wave amplitudes provide insight into hydroacoustic blockage, which is essential for understanding detection capability. Seismic body-wave conversions beneath the arrays can be timed to 1 second or better, and provide important location constraints for subocean events occurring far from continental stations.

9:45

2aAO6. Directionality of acoustic T-phase signals from shallow submarine earthquakes. Ross Chapman (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, chapman@uvic.ca)

Acoustic transients radiated from shallow undersea earthquakes were studied in an experiment carried out using a towed horizontal line array operating in the South Fiji Basin. The transient signals consisted of P, S, and T phases, with the T-phase signal from each earthquake lasting for about 5 min. During this period, the directionality of the T-phase signal was determined by processing the array data with a conventional beamformer. The weakest part of the signal arrived first on a direct bearing between the earthquake source and the array. However, subsequent stronger components of the T phase arrived from different directions farther south of the source, in a region where a number of seamounts rose within the sound channel. A simple model based on ray-path travel times for p-wave travel in the earth and in the water suggests that the later components of the T-phase signal are radiated into the water by downslope propagation from the seamounts and ridges. The initial weaker components may be scattered into the sound channel by leakage from the P and S phases relatively closer to the array.

10:05–10:20 Break

Contributed Papers

10:20

2aAO7. Modal scattering and T-waves: Sediment amplification and source effects. Robert I. Odom (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105) and Darin J. Soukup (Univ. of Washington, Seattle, WA 98195)

If the Earth were a plane-layered semi-infinite half-space or a radially symmetric sphere oceanic, T-waves could not exist. This is apparent because the source depth of T-wave producing earthquakes is greater than the depth at which the low order modes comprising the T-waves have any significant amplitude. Bottom roughness provides a mechanism by which energy from high order source modes can be scattered into the lower order modes. The efficiency of this scattering is improved when there is a layer of sediments overlying the higher speed upper ocean crust. Some of the source modes develop a large anti-node, reminiscent of a Scholte wave, at the water sediment interface. Bottom roughness serves as a sheet of secondary sources placed directly on the anti-node of these modes, and contributes a significant amount of energy to the T-waves. However, a significant fraction of the total T-wave energy is also provided by small scattered contributions from many higher modes of the continuum spectrum, which is modeled using the locked mode approximation. Source mechanism modeling of T-wave excitation shows that normal fault earthquakes are inefficient generators of T-waves.

10:35

2aAO8. Coupling of acoustic normal modes and seismo-acoustic surface waves in variable-depth shallow water. Oleg A. Godin (CIRES, Univ. of Colorado and NOAA/Environ. Technol. Lab., Mail Code R/ET1, 325 Broadway, Boulder, CO 80305)

Marine sediments support Rayleigh-type surface waves which can propagate along the sea floor in arbitrarily shallow water and even on shore. At frequencies above a few Hertz, because of a strong attenuation of compression and especially shear waves in the sediments, the surface

waves can significantly contribute to the acoustic field far from the shore only through their coupling to volume waves in the water. We study theoretically excitation of acoustic normal modes by the seismo-acoustic surface waves in a shallow-water waveguide with a sloping bottom. The role of bottom stratification is assessed by comparing acoustic fields generated by the Rayleigh wave, which is supported by a homogeneous elastic bottom, and by the fundamental mode of surface waves, which are supported by the bottom with a linear profile of shear rigidity. It is found that the coupling primarily occurs in the vicinity of a modal cutoff. The first normal mode is the one most strongly excited by the surface waves. For realistic values of geoacoustic parameters, the energy conversion rate can reach tens of percent, making the surface-to-volume wave conversion an efficient mechanism of the oceanic waveguide excitation by acoustic sources on shore or in very shallow water.

10:50

2aAO9. Listening to distant earthquakes with hydrophones on autonomous oceanic floats. Frederik Simons, Guust Nolet (Princeton Geosciences, Guyot Hall, Princeton, NJ 08540, fjsimons@princeton.edu), and Jeff Babcock (UCSD, La Jolla, CA 92093)

We are mounting a hydrophone on autonomous floats capable of drifting below the sound channel and surfacing to communicate data by a satellite link. Using new, intelligent algorithms for the automatic identification and discrimination of seismic phases, we recognize teleseismic arrivals in the presence of local *P*, *S*, and *T* phases, ship and whale noise, and other contaminating factors such as airgun surveys. Our approach combines time-domain methods with spectrogram analysis and with wavelet methods. To maximize battery life, we optimize the efficiency of our algorithms and their numerical implementation. Our algorithms were tested on data from tethered hydrophones from two arrays anchored to the Mid-Atlantic Ridge and the East Pacific Rise. Our prototype device was succesfully tested in a dive to 700 m off the coast of La Jolla. We acquired a valuable 31-h data stream suitable to characterize the ambient noise and were able to identify a number of engineering problems with the hydrophone sensitivity. We expect detecting teleseisms with magnitudes superior to 6.3, and thus adding at least 200 high-quality recordings to the global catalog over the life span of a single instrument—well worth the $15 000 manufacturing price and negligible deployment costs on ships of opportunity.

11:05

2aAO10. Hydroacoustic events located near the Atlantis (30N) and Kane (2330N) transform faults on the MAR. Clare M. Williams, Ralph A. Stephen, and Deborah K. Smith (WHOI, 360 Woods Hole Rd. (MS 24), Woods Hole, MA 02543-1542, rstephen@whoi.edu)

Hydroacoustic arrivals detected by a hydrophone array on the Northern Mid-Atlantic Ridge (MAR) are used to better understand *T* phases from oceanic crust earthquakes. *T*-phase events are selected from two study areas: the inside corner and ridge transform intersection at the eastern ends of the Kane (2330N) and Atlantis (30N) transform faults. Both are regions of high relief (5000 m) and events are located throughout the massif and transform valleys. We investigate the spatial distribution of *T*-phase locations as a function of water depth. We examine the characteristics of *T* phases as a function of both distance from the event to each hydrophone and event location water depth. Finally, we use a ray-trace model to test for bathymetric blockage along the propagation path of the *T* phase. We observe that acoustic magnitudes of *T* phases show no dependence on water depth of event location. This is opposite to predictions from current *T*-phase generation models which show water depth dependence. There may be a correlation between acoustic magnitude and propagation path topography, but results are sensitive to the sound velocity profile used in the model. Our results underscore the complexity of *T*-phase generation and propagation.

TUESDAY MORNING, 25 MAY 2004 NEW YORK BALLROOM A, 7:55 A.M. TO 12:00 NOON

Session 2aBB

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: High Intensity Focused Ultrasound I

Cheri Deng, Cochair
Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106

Shahram Vaezy, Cochair
University of Washington, HSB Box 356540, Seattle, Washington 98195-6540

Chair's Introduction—7:55

Invited Papers

8:00

2aBB1. The feasibility of noninvasive image-guided treatments of brain disorders by focused ultrasound. Kullervo Hynynen, Gregory Clement, Nathan McDannold, Natalia Vykhodtseva (Brigham and Women's Hospital and Harvard Med. School, 75 Francis St., Boston, MA 02115, kullervo@bwh.harvard.edu), P. Jason White, Ferenc A. Jolesz, and Nickolai A. Sheikov (Brigham and Women's Hospital and Harvard Med. School, Boston, MA 02115)

Brain disorders, such as tumors, functional problems, etc., are difficult to treat and the invasive interventions often disturb surrounding brain tissue, resulting in complications. In addition, the delivery of therapeutic agents to the brain via the blood supply is often impossible because the blood brain barrier protects the brain tissue from foreign molecules. Our hypothesis has been that transcranial therapeutic ultrasound exposures can be delivered with an optimized phased-array system. We have demonstrated that highly focused therapeutic ultrasound beams can be accurately delivered through an intact human skull noninvasively. Furthermore, we demonstrated using *ex vivo* human skulls that we can use CT-derived information to predict the phase shifts required for correcting

the wave distortion. We have also developed a method to focally disrupt the blood brain barrier without damaging the neurons in the targeted tissue volume. This may allow delivery of therapeutic or diagnostic agents into image-specified locations. Successful transcranial delivery of ultrasound in a clinical setting may have a major impact on the treatment of brain disorders in the future.

8:20

2aBB2. Design and testing of an ultrasonic system with integrated diagnostic and therapy modes. Frederic L. Lizzi (Riverside Res. Inst., 156 William St., New York, NY 10038-2609)

Our laboratories developed and tested a system integrating high-intensity focused ultrasound (HIFU) with diagnostic ultrasound procedures for aiming, delivering, and monitoring treatment exposures. The system employs 5-MHz, 5-element spherical-cap HIFU arrays with annular electrodes (for variable focusing) or strip electrodes (for inducing broad lesions of controlled width). Central apertures in these arrays house a coaxial diagnostic phased array (48 elements), interfaced with a custom digital system for visualization and RF echo acquisition. Several operational modes were investigated. A preexposure mode permits the HIFU focal point to be aimed at the desired location in B-mode images. It also assesses harmonic generation in the HIFU beam, using the diagnostic array to receive and spectrally analyze echoes from a short HIFU pulse aimed at targeted tissues. Radiation-force elastography uses the diagnostic array to monitor tissue displacements induced by brief exposures from the HIFU transducer; pre- and posttherapy data are compared to delineate mechanical alterations in lesioned volumes. Acquired pre- and post-RF data are also spectrally analyzed to map changes in scattering and attenuation indicative of lesion production. These modes have been successfully tested under *in vitro* conditions and are now entering use in animal studies aimed at treating cancer and cardiac diseases.

8:40

2aBB3. The physical acoustics of HIFU exposure: A suite of *in vitro* studies. Ronald A. Roy, R. Glynn Holt, and Robin O. Cleveland (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ronroy@bu.edu)

Exposing tissue and blood media to high-intensity focused ultrasound leads to a broad range of physical and biological effects. In addition to tissue heating from thermo-viscous absorption, one encounters nonlinear and second order phenomena such as wave-front shocking, radiation stress, acoustic streaming, bubble nucleation and cavitation. All of these processes impact one's ability to effect controlled tissue heating. This paper reviews a suite of studies under way at Boston University that focus on developing a quantitative understanding of the physical acoustics of high-intensity focused ultrasound. We describe research into HIFU heating and both uniform and flow-through tissue phantom and address the potentially critical roles played by bubbles and acoustic cavitation. [Work supported by the U.S. Army and the Center for Subsurface Sensing and Imaging Systems via NSF ERC Award No. EEC-9986821.]

9:00

2aBB4. Approaches to overcome current limitations of HIFU treatment. Shin-ichiro Umemura, Ken-ichi Kawabata, Kazuaki Sasaki, Takashi Azuma (Central Res. Lab., Hitachi Ltd., Kokubunji, Tokyo 185-8601, Japan, sumemura@crl.hitachi.co.jp), Kazunari Ishida, Jun Kubota (Hitachi Medical Corp., Chiba 277-0804, Japan), Mitsuyoshi Ichihara, and Takashi Okai (Showa Univ. School of Medicine, Tokyo 142-8666, Japan)

Noninvasive therapy with HIFU has been successfully applied to transrectal treatment of prostate cancer as well as benign prostate hyperplasia. However, there are two major technical reasons why its clinical application to other organs is currently limited: (1) low throughput of treatment and (2) lack of penetration to deep tissues. To multiply the throughput, a split-focus technique, in which the focal spot is enlarged primarily in the lateral direction, was developed. An electronically variable focus array transducer was also developed to enhance the throughput. An approach to treat a large volume of uterus myoma by coagulating its feeding arteries has been studied. The tissue volume to be coagulated can be thereby reduced by orders of magnitude. The penetration and throughput can potentially be improved at the same time by delivering a microbubble agent to the target tissue. It was theoretically predicted that a microbubble agent could multiply the ultrasonic tissue absorption. The effectiveness of this approach was confirmed in animal experiments using Optison. Real-time monitoring of tissue coagulation during HIFU exposure also can enhance the throughput through preventing excess deposition of ultrasonic energy. Monitoring coagulation by imaging local displacement in tissue with ultrasound will be discussed as well.

9:20

2aBB5. Thermal ablation by interstitial ultrasound. Cyril Lafon, David Melodelima, Jean-Yves Chapelon, and Dominique Cathignol (INSERM U556, 151 Cours Albert Thomas, 69003 Lyon, France)

Interstitial ultrasound applicators have been proposed for treating deep-seated tumors that are unreachable with extracorporeal high intensity focused ultrasound. The technique consists of bringing the ultrasound source as close as possible to the target in order to minimize the effects of attenuation and phase aberration along the ultrasound pathway. We have designed interstitial applicators with flat ultrasound transducers that operate at a frequency between 3.5 and 20 MHz. Nonfocused transducers require working at high frequencies to obtain coagulation in a short period of time. Flat transducers were chosen because the pressure drop in the near field is only due to absorption, and extended therapeutic depth can be achieved. A rotation of the transducer results in a cylindrical or sector-based volume of necrosis. To avoid these mechanical displacements, a cylindrical array has been manufactured; it allows for electronic rotation of a plane wave. Imaging modalities like endoscopic ultrasound and MRI were tested for guiding our treatments. Clinical trials are now being performed for the thermal ablation of digestive tumors and preliminary results are extremely promising.

9:40

2aBB6. Comparison between coupled KZK-BHTE numerical simulations and scanned HIFU exposures in excised bovine liver. Marilee A. Andrew, Andrew A. Brayman, Peter J. Kaczkowski, and Steven G. Kargl (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698)

The use of moving high intensity focused ultrasound (HIFU) treatment protocols is of interest in achieving efficient formation of large-volume lesions in tissue. However, potentially unwanted thermal effects, such as prefocal heating, should be considered. A KZK acoustic model coupled with the BioHeat Transfer Equation has been extended to simulate multiple, moving scans in tissue. Simulation results are compared with experimental data collected over a range of exposure regimes for linear and concentric circular scans with a 3.5-MHz single-element transducer in *ex vivo* bovine liver. Of particular interest are investigating prefocal thermal buildup and ablating the central core of a circular pattern through conductive heating, that is without direct HIFU exposure. Qualitative agreement is observed between experimental and simulated data; limits of the predictive capability of the model in cavitation regimes will be discussed. [Support provided by the U.S. Army Medical Research Acquisition Activity through The University of Mississippi under terms of Agreement No. DAMD17-02-2-0014. The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of U.S. Army Medical Research Acquisition Activity or The University of Mississippi.]

9:55

2aBB7. Cavitation detection during and following HIFU exposure *in vitro*. Constantin-C. Coussios (Dept. of Eng. Sci., Univ. of Oxford, Parks Rd., Oxford OX1 3PJ, UK, constantin.coussios@eng.ox.ac.uk), Caleb H. Farny, Charles R. Thomas, Robin O. Cleveland, R. Glynn Holt, and Ronald A. Roy (Boston Univ., Boston, MA 02215)

Is the existence of a sustained hyperechogenic region on B-scan images following HIFU treatment a necessary and sufficient condition for cavitation to have occurred during HIFU exposure? Three means of cavitation monitoring were used synchronously, before, during, and after continuous-wave HIFU exposure of an agar-graphite tissue phantom. A 1.1-MHz HIFU transducer was confocally aligned with a 15-MHz passive cavitation detection (PCD) transducer and a 5-MHz scan head. The HIFU pressure amplitude was increased in steps of 0.26 MPa every 5 s. A peak detector recorded the peak PCD signal level and a dynamic signal analyzer monitored broadband noise emissions (5–10 MHz). The sudden onset of a PCD output signal occurred at a peak-negative focal pressure of 1.25 MPa; no post-HIFU hyperechogenic region was visible on the B-scan images. A hyperechogenic region did eventually appear, but only for focal pressures in excess of 1.8 MPa. Inertial cavitation can therefore occur during HIFU exposure in the absence of a post-HIFU hyperechogenic region. The focal pressure for which such a region is observed can be as much as 50% higher than the threshold pressure for inertial cavitation inception. [Work supported by the ASA and the NSF Center for Subsurface Sensing and Imaging Systems.]

10:10–10:30 Break

10:30

2aBB8. Fluorescent real-time monitoring of HIFU cardiac focal ablation. Fujian Qu, Vladimir Nikolski, Igor Efimov, and Cheri Deng (Dept. of Biomed. Eng., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106-7207, cxd54@cwru.edu)

To study HIFU cardiac ablation, florescent imaging was used to monitor in real time the electrophysiology changes of cardiac tissues during focal HIFU ablation. We applied HIFU ablation of AV nodal and ventricular preparations of Langendorff-perfused rabbit heart while monitoring electrical activity in real-time. HIFU energy was applied to ablate the AV node and ventricular tissue of Langendorff-perfused rabbit hearts while monitoring electrical activity in real-time with fluorescent voltage-sensitive dye imaging and surface electrodes. HIFU was generated using a spherical piezoelectric ceramics transducer (diameter 42 mm, F-number 1.2) at 4.23 MHz. When HIFU was applied to ventricular epicardium fluorescent imaging it revealed gradual reduction of the plateau phase and amplitude of the action potential. Subsequently conduction block and cell death were observed at the site of ablation. In our study HIFU produced focal lesions of 0.2–0.8 mm for 10–60-s applications. When HIFU was applied to the AV node, fluorescent imaging and electrograms revealed the development of the AV block.

10:45

2aBB9. Feasibility of *in vivo* cardiac HIFU ablation. Kana Fujikura, Ryo Otsuka, Shunichi Homma (Columbia Presbyterian Med. Ctr. PH3-133, 630 W. 168th St., New York, NY 10032, kf2113@columbia.edu), Robert Muratore, Jeffrey A. Ketterling, and Frederic L. Lizzi (Riverside Res. Inst., New York, NY 10038-2609)

The potential for cardiac applications of HIFU remains unexamined. In order to create reproducible lesions in a beating heart, it is necessary to maintain focusing at a certain position within moving myocardial tissue. One technique is to use multiple short HIFU exposures (0.2 s) and synchronize them with an EKG signal and respiration. The left ventricular free wall (LVFW) of calf hearts were cut into 4-cm cubes, degassed in phosphate buffer saline (PBS), and heated to 37 °C. An 80-mm-diam spherical-cap transducer with a focus of 90 mm was operated at a frequency of 4.67 MHz and a nominal focal point intensity of 26.9 kW/cm^2. The transducer was coupled to the LVFW using degassed PBS. First, the effect of pericardial fat, focal depth, and temperature on lesion size was individually evaluated. Then the 0.2-s HIFU exposure was applied 10 to 30 times at 4-s intervals. The same HIFU transducer was applied to an open-chest canine LVFW with the same triggering protocol. Dimensions of all lesions were measured by visual examination of the fresh, unstained tissue. A histopathological examination of the lesion was also performed. The *in vivo* lesions were created in similar size to those *in vitro*.

11:00

2aBB10. Experimental high-intensity focused ultrasound lesion formation in cardiac tissue. Robert Muratore, Andrew Kalisz, Paul Lee, Frederic Lizzi (Riverside Res. Inst., 156 William St., New York, NY 10038-2609, muratore@rrinyc.org), Kana Fujikura, Ryo Otsuka, and Shunichi Homma (Columbia Univ. College of Physicians and Surgeons, New York, NY 10032)

High-intensity focused ultrasound (HIFU) (4.5–7.5 MHz) was used to form lesions in cardiac tissue, with an ultimate objective of treating conditions such as hypertrophic cardiomyopathy and ventricular tachycardia. Ultrasound attenuation coefficients were experimentally determined *in vitro* for calf myocardial tissue, both muscle and pericardial fat. These coefficients were employed in computational models of linear beam propagation, tissue heating profiles and thermal lesion formation for a variety of focused transducers. Modeling was performed for continuous and pulsed exposures. These models suggested initial power levels and exposure durations for *in vitro* experiments on calf ventricles and septa and *ex vivo* experiments on canine whole hearts. Repeatability of lesion size and placement was studied as power and exposure parameters varied around the initial values. With these experimental results, power and exposure parameters were selected to create lesions *in vivo* in canine ventricles and septa in open-chest, anesthetized dogs. Pulsed exposures were synchronized to cardiac and respiration cycles to ensure accurate placement of the lesions. These initial *in vivo* experiments showed that HIFU treatments in the beating heart are feasible; they also identified refinements that are now being implemented for better control of lesion size and placement. [Work supported by NCI and NHLBI Grant 5R01 CA84588.]

11:15

2aBB11. The relative effects of cavitation and nonlinear ultrasound propagation on HIFU lesion dynamics in a tissue phantom. Vera A. Khokhlova (Dept. of Acoust., Faculty of Phys., M. V. Lomonosov Moscow State Univ., Leninskie gory, Moscow 119992, Russia, vera@acs366.phys.msu.su), Michael R. Bailey, Justin Reed, and Peter J. Kaczkowski (Univ. of Washington, Seattle, WA 98105)

The relative importance of the effects of acoustic nonlinearity and cavitation in HIFU lesion production is studied experimentally and theoretically in a polyacrylamide gel. A 2-MHz transducer of 40-mm diameter and 45-mm focal length was operated at different regimes of power, and in cw or duty-cycle regimes with equal mean intensity. Elevated static pressure was applied to suppress bubbles, increase boiling temperature, and thus to isolate the effect of acoustic nonlinearity in the enhancement of lesion production. Experimental data were compared with the results of simulations performed using a KZK acoustic model combined with the bioheat equation and thermal dose formulation. Boiling and the typical tadpole-shaped lesion shifting towards the transducer were observed under standard atmospheric pressure. No boiling was detected and a symmetric thermal lesion formed in the case of overpressure. A delay in lesion inception time was registered with overpressure, which was hypothesized to be due to suppressed microbubble dynamics. The effect of acoustic nonlinearity was revealed as a substantial decrease in the lesion inception time and an increase in the lesion size for high-amplitude waves under both standard and overpressure conditions. [Work supported by ONRIFO, NASA/NSBRI, NIH Fogarty, and CRDF grants.]

11:30

2aBB12. Effect of nonlinearity on lesion formation for high-intensity focused ultrasound (HIFU) exposures. Paul Lee, Frederic L. Lizzi, Jeffrey A. Ketterling (Riverside Res. Inst., 156 William St., New York, NY 10038, lee@rrinyc.org), and Christopher J. Vecchio (Spectrasonics Imaging, Inc., Wayne, PA 19087)

This study examined the effects of nonlinear propagation phenomena on two types of HIFU transducers (5 MHz) being used for thermal treatments of disease. The first transducer is a 5-element annular array. The second is a transducer with a 5-strip electrode; its multilobed focused beam is designed to efficiently produce broad, paddle-shaped lesions. The beam patterns of these transducers were computed using a variety of excitation patterns for electronic focusing of the annular array and variation of lesion size for the strip-electrode transducer. A range of intensities was studied to determine how nonlinear propagation affects the beam shape, constituent frequency content, grating lobes, etc. These 3D computations used a finite-amplitude beam propagation model that combined the angular spectrum method and Burger's equation to compute the diffraction and nonlinear effects, respectively. Computed beam patterns were compared with hydrophone measurements for each transducer. The linear and nonlinear beam patterns were used to compute the absorbed thermal dose, and the bioheat equation was evaluated to calculate 3D temperature rises and geometry of induced lesions. Computed lesion sizes and shapes were compared to *in vitro* lesions created by each HIFU transducer. [Work supported by NCI and NHLBI Grant 5R01 CA84588.]

11:45

2aBB13. Experimental verification of enhancement of HIFU-induced heating of tissue mimicking phantoms due to acoustic nonlinearity. Tatiana V. Sinilo, Vera A. Khokhlova, and Oleg A. Sapozhnikov (Dept. of Phys., Moscow State Univ., Leninskie Gory, Moscow 119992, Russia, tanya@acs366.phys.msu.ru)

A possibility of enhancement of heat deposition of high-intensity focused ultrasound (HIFU) in tissue mimicking phantoms due to acoustic nonlinearity is experimentally studied. A gelatin sample of high concentration was used to model medium with acoustic properties similar to those of biological tissue. Several sets of pulse-periodic regimes with the same mean power but different pulse amplitudes and durations were used. In regimes with higher amplitudes (lower duty cycles), the waveform in the focal region was more strongly distorted because of acoustic nonlinearity. In the linear medium all the regimes of one set would give the same heating. Optical shadow pictures of the HIFU-heated region were taken in the regimes with different duty cycles but the same mean power. The heated region size increased for lower duty cycles, but the shadow shape remained regularly cigar-like in all the regimes, even when the ultrasound waveform was shocked. The transmitted signal measured by a hydrophone was stable. This showed that cavitation was not pronounced. Thermocouple measurements and theoretical modeling of gelatin temperature in the focal region of the ultrasound beam showed significant increase of HIFU heat deposition in the presence of shocks in the waveform. [Work supported by CRDF and RFBR.]

TUESDAY MORNING, 25 MAY 2004 CONFERENCE ROOM K, 8:00 TO 10:15 A.M.

Session 2aMUa

Musical Acoustics: Measurement and Analysis

James P. Cottingham, Chair
Department of Physics, Coe College, Cedar Rapids, Iowa 52402

Contributed Papers

8:00

2aMUa1. Resonant and nonresonant sound radiation from a guitar. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

When forces are applied to a flexible structure, a worker in structural acoustics who is concerned with the radiated sound will normally be interested in how much of the sound is arising from the resonant response of the structure compared to the nonresonant response. Lightweight panels below their critical frequency are likely to have a significant amount of sound arising from nonresonant response due to mass inertial response localized around the location of force application. Since a guitar top is such a structure, it is natural to inquire about the comparison between resonant and nonresonant radiation from a guitar over its normal range of frequencies. Measurements have been made on a student Yamaha and a concert quality Modranas/Connor guitar. The results of these measurements and their interpretation will be discussed.

8:15

2aMUa2. Noncontact modal analysis of a pipe organ reed using airborne ultrasound stimulated vibrometry. Thomas M. Huber (Dept. of Phys., Gustavus Adolphus College, 800 College Ave., Saint Peter, MN 56082, huber@gustavus.edu), Mostafa Fatemi, Randall R. Kinnick, and James F. Greenleaf (Mayo Clinic College of Medicine, Rochester, MN 55905)

The goal of this experiment was to excite and measure, in a noncontact manner, the vibrational modes of the reed from a reed organ pipe. To perform ultrasound stimulated excitation, two ultrasound beams in air of different frequencies were directed at the reed; the audio-range beat frequency between these ultrasound beams induced vibrations. The resulting vibrational deflection shapes were measured with a scanning vibrometer. The modes of any relatively small object can be studied in air using this technique. For a 36 mm by 7 mm clamped brass reed cantilever, displacements and velocites of 5 μ and 4 mm/s could be imparted at the fundamental frequency of 145 Hz. Using the same ultrasound transducer, excitation across the entire range of audio frequencies was obtained, which was not possible using audio excitation with a speaker. Since the beam was focused on the reed, ultrasound stimulated excitation eliminated background effects observed during mechanical shaker excitation, such as vibrations of clamps and supports. We will discuss the results obtained using single, dual, and confocal ultrasound transducers in AM and unmodulated CW modes, along with results obtained using a mechanical shaker and audio excitation using a speaker.

8:30

2aMUa3. Measurement of sound intensity near flue organ pipes. Matthias Scholz, Essa Ibrahim Abdugalder (Chalmers Rm. Acoust. Group, Appl. Acoust., Chalmers, Gothenburg SE-41296, Sweden), Mendel Kleiner (Rensselaer Polytech. Inst., Troy, NY 12180), and Munetaka Yokota (Gothenburg Univ., Gothenburg, Sweden)

Organ voicers sometime experience that organ pipes, within an organ case, acoustically couple to one another. The coupling can be not only between a pipe and its mirror image in the case wall but also between adjacent pipes tuned to the same frequency. This paper reports on intensity measurements done to see the influence of a nearby pipe on the intensity field from a sounding pipe. The sounding pipe was driven pneumatically, and the sound intensity was measured in a plane along the length axis of the pipe. The secondary pipe was positioned, in parallel at different distances, alongside the sounding pipe. The results indicate only a fairly small influence by the secondary pipe on the intensity map of the sounding pipe. [Work supported by the Chalmers Foundation.]

8:45

2aMUa4. The application of cochlear audio analysis techniques to electro-acoustic music. John A. Mills III (Elec. and Computer Eng. Dept., Univ. of Texas, 1 University Station 0803, Austin, TX 78712, nodog@mail.utexas.edu)

Electro-acoustic music has been ignored by music theorists for years. One theory suggests that this deficiency is due to the lack of an objective visual representation of this type of music. This research is focused upon the audio analysis involved in the creation of a ''pseudoscore'' for electro-acoustic music. Computer analysis of music is a complicated task. The ''holy grail'' of automatic computer music analysis has most often been the translation of a continuous pressure variation into traditional Western musical notation. Since electro-acoustic music is rarely able to be transcribed into this type of notation, the automatic analysis of electro-acoustic music confounds many previous algorithms. A top-down approach is suggested in order to extract acoustic and musical information from recordings of electro-acoustics music. A top-down approach has already proven successful in extraction of musical tempo [E. Scheirer, M.I.T. Ph.D. thesis (2000)]. Since humans are able to translate a continuous pressure variation into useful acoustic information, when using a top-down approach to automatic music analysis, a cochlear model is used as a front end. Although this model apparently encodes some redundant information, cross- and auto-correlation techniques allow easier extraction of some acoustic information. This presentation will detail the current state of this research.

9:00

2aMUa5. Tuning and rotating nearly degenerate modes in handbells. John R. Buschert (Goshen College, 1700 S. Main, Goshen, IN 46526) and Jennifer L. Springer (The Univ. of California, Davis, Davis, CA 95616)

We have studied nearly degenerate pairs of vibrational modes in a C4 handbell. Manufacturers seek to eliminate the beating of these modes by careful choice of the clapper position but this can't always be done satisfactorily. Using electronic holography, we are studying the effects of small, localized mass changes on the frequency and orientation of these mode pairs. One expects one mode of the pair to rotate until it has an anti-node at the point of the added mass and the other to rotate until it has a node there. Then adding further mass should only change the frequency of the first mode. We find some mode pairs respond in just the expected manner but others behave more strangely. [Work supported by an NSF MRI grant.]

9:15

2aMUa6. Low-cost coding of directivity information for the recording of musical instruments. Jonas Braasch, William L. Martens, and Wieslaw Woszczyk (Faculty of Music, McGill Univ., 555 Sherbrooke St. W., Montreal, QC H3A 1B9, Canada, braasch@music.mcgill.ca)

Most musical instruments radiate sound according to characteristic spatial directivity patterns. These patterns are usually not only strongly frequency dependent, but also time-variant functions of various parameters of the instrument, such as pitch and the playing technique applied (e.g., plucking versus bowing of string instruments). To capture the directivity information when recording an instrument, Warusfel and Misdariis (2001) proposed to record an instrument using four channels, one for the monopole and the others for three orthogonal dipole parts. In the new recording setup presented here, it is proposed to store one channel at a high sampling frequency, along with directivity information that is updated only every few milliseconds. Taking the binaural sluggishness of the human auditory system into account in this way provides a low-cost coding scheme for subsequent reproduction of time-variant directivity patterns.

9:30

2aMUa7. Acoustical studies of the American reed organ. James P. Cottingham (Phys. Dept., Coe College, Cedar Rapids, IA 52402)

The reed organ enjoyed a period of great popularity in North America which reached a peak in the late 19th century, when thousands of instruments per year were manufactured and sold in the United States and Canada. Displaced by the emergence of the upright piano, the reed organ had very much fallen out of favor by 1929. In the past decade a number of acoustical investigations have been undertaken on the instrument known as the American reed organ. Observations of reed motion and velocity have been made with electronic proximity sensors and a laser vibrometer system. The variation of the frequency and amplitude of reed vibration as a function of blowing pressure has been explored in some detail and the results compared with predictions of a simple theoretical model. Measurements have been made of the spectrum of the near-field sound including the effects of changes in dimensions of the reed cell. While most treatments of free reed oscillation approximate the reed vibration as a sinusoidal oscillation of a cantilever beam in the fundamental transverse mode, recently some evidence of higher transverse modes and torsional modes of vibration have been observed.

9:45

2aMUa8. Calculation of impulse responses with a cellular automata algorithm. Ana Barjau (Dept. of Mech. Eng., Polytech. Univ. of Catalunya, Diagonal 647, 08028 Barcelona, Spain, ana.barjau@upc.es)

The air columns in musical instruments usually have a predominant dimension and thus are very often modeled as 1D systems where uniparametric waves propagate. Different algorithms can be found in the literature to simulate this propagation. The more widely used are finite difference schemes and delay lines. A finite difference scheme (FD) is a numerical integration of a differential formulation (the wave equation), while delay lines (DL) use analytical exact solutions of the wave equation over finite lengths. A new and different approach is that of a cellular automaton (CA) scheme. The underlying philosophy is opposite those of FD and DL, as the starting point is not the wave equation. In a CA approach, the phenomenon to be studied is reduced to a few simple physical laws that are applied to a set of cells representing the physical system (in the present case, the propagation medium). In this paper, a CA will be proposed to obtain the impulse response of different bore geometries. The results will be compared to those obtained with other algorithms.

10:00

2aMUa9. Transformations and invariances in equally tempered scales. Alpar Sevgen (Dept. of Phys., Boğaziçi Univ., Bebek 34342, İstanbul, Turkey)

In this work the invariance of the number of accidentals in equally tempered scales (ETSs) under various transformations are proven *analytically*. ETSs with N semitones, M notes and interval structures $\mathbf{n}=\{n_1,n_2,\ldots,n_M\}$, where n_i is the number of semitones between adjacent notes, form multiplets of N scales. Member scales of a multiplet are ordered and indexed by scale labels $c_k=kM$ *reduced modulo* (N), where $k=0,1,\ldots,N-1$. Ordered members of the multiplets having the interval structures $\{\mathcal{C}^j\mathbf{n}\}$ and $\{\mathcal{C}^\ell\mathcal{MI}\mathbf{n}\}$ (where $\mathcal{C}$ is a cyclical permutation, $j,\ell=0,1,\ldots,M-1$, and $\mathcal{MI}$ is a mirror image or flip operation on $\mathbf{n}$) are shown to have the *same* number of sharps and the *same* number of flats, $\mathcal{N}_\sigma(c_k,\mathcal{C}^j\mathbf{n})=\mathcal{N}_\sigma(c_k,\mathcal{C}^\ell\mathcal{MI}\mathbf{n})$ (where $\sigma=\sharp$ or $\flat$, for all j, ℓ and for any scale label c_k). Complementary scales $\mathbf{n}_C$ related to $\mathbf{n}$ by a note $\leftrightarrow$ no-note transformation (e.g., major $\leftrightarrow$ pentatonic), however, have scale labels with opposite signs, and the number of sharps and flats are interchanged $\mathcal{N}_\sharp(c_k,\mathbf{n})=\mathcal{N}_\flat(-c_k,\mathbf{n}_C)$. This explains the degeneracy of the complexity parameter defined as the sum of the number of sharps and flats, proposed as a measure for ETS, explored numerically and reported earlier.

TUESDAY MORNING, 25 MAY 2004 CONFERENCE ROOM K, 10:30 TO 11:55 A.M.

Session 2aMUb

Musical Acoustics: Musical Instruments Developed During the ASA Era

Paul A. Wheeler, Cochair
1595 North 1600 East, Logan, Utah 84341

Ian M. Lindevald, Cochair
Science Division, Truman State University, Kirksville, Missouri 63501

Chair's Introduction—10:30

Invited Papers

10:35

2aMUb1. Musical instrument technology of the 20th century. Paul Wheeler (Utah State Univ., 4120 Old Main Hill, Logan, UT 84322, paul.wheeler@ece.usu.edu)

This paper presents a brief history of the technical development of musical instruments during the 20th century. Starting with early electronic instruments (such as the Theremin—1917) invented prior to the organization of ASA, the history includes the development of electronic organs, synthesizers, and computer music. This paper provides an introduction to the session, giving a framework for the papers which follow in the session.

10:55

2aMUb2. The Electronic Valve Instrument (EVI), an electronic musical wind controller for playing synthesizers. Nyle A. Steiner (269 E. 300 South, Provo, UT 84606, lenyr@earthlink.net)

The Electronic Valve Instrument (EVI) is an electronic musical wind instrument with playing techniques similar to that of a trumpet. Invented by Nyle Steiner in the early 1970's, it was designed to give the performer control of dynamics from breath pressure and the ability to make a humanly generated vibrato. Other musical paramaters can be controlled as well. It has a playing range of seven octaves (similar to that of a piano). When musical lines are played using this instrument (controller) connected to an electronic music synthesizer, the sound is much more natural sounding and expressive than when a normal musical keyboard is used. The evolution of this instrument from the pre-Midi era to it latest Midi configuration, principles of operation, synthesizer programming, and its wide use in movie and TV scoring will be discussed. The EVI has played featured musical lines in many major movie soundtracks and TV shows such as *Apocalypse Now, Witness, Dead Poets Society, Fatal Attraction, No Way Out, Gorillas in the Mist*, and many others. The EVI design has also been adapted as an Electronic Woodwind Instrument (EWI) by Nyle Steiner and has been manufactured and sold worldwide by the AKAI Co. in Japan.

11:15

2aMUb3. New paradigms for musical control—A decade of development at the MIT Media Lab. Joseph A. Paradiso (MIT Media Lab, E15-327, 20 Ames St., Cambridge, MA 02139, joep@media.mit.edu)

As electronic musical instruments liberate the action and energy of control from physical sound production, they are free to mutate into many different forms—the constraints on instrument design have shifted from physics to ergonomics, applications, and aesthetics. Low-cost sensors enable stimuli of all types to act as input, and with a computer interposed between action and sound production, essentially any sonic or musical dynamic can be mapped onto any gesture or activity with an increasingly high degree of interpretation or ''mapping,'' Accordingly, the notion of a musical instrument is being redefined, and as possibilities broaden, some researchers and artists are striving to break boundaries while others work to quantify and understand expanded metrics for musical interaction. Over the past decade, the author and his colleagues have adapted a wealth of sensor technologies and developed many interaction paradigms to scratch away at the evolving frontier of electronic musical instruments [J. Paradiso, ''Electronic music interfaces: new ways to play,'' IEEE Spectrum **34**(12), 18–30 (1997)]. This presentation will review the status of electronic music controllers, provide a snapshot of current issues that the field is facing, and present various examples of new musical interfaces developed at the MIT Media Lab.

11:35

2aMUb4. Exploring a Van der Pol woodwind. Ana Barjau and Meritxell Genesc (Dept. of Mech. Eng., Polytech. Univ. of Catalunya, Diagonal 647, 08028 Barcelona, Spain, ana.barjau@upc.es)

The behavior of woodwinds can be studied from a theoretical point of view with a variety of mathematical models, ranging from very realistic to fairly simple ones. All of them have in common a nonlinearity responsible for the self-sustained oscillation. In simple models, this nonlinearity is represented through an algebraic equation, while in realistic ones it may be split into different differential equations representing the dynamics of the mechanical device (reed, lips . . .) and of the air flowing through it. In this paper, an intermediate model is explored. A Van der Pol oscillator is used to represent just the nonlinear device of a woodwind (reed valve, lips . . .), while the air column is simplified into a cylindrical or conical one, and its behavior is described through an impulse response. The global system is thus represented mathematically through two coupled differential equations, a linear one containing delayed variables and a nonlinear one. Their integration leads to different sounds, some of which are close to those produced by real instruments.

TUESDAY MORNING, 25 MAY 2004 VERSAILLES TERRACE, 8:35 TO 11:40 A.M.

Session 2aNS

Noise and Psychological and Physiological Acoustics: Noise Impact Evaluation: Old and New I

Brigitte Schulte-Fortkamp, Cochair
Institute of Technical Acoustics, Technical University Berlin, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany

Klaus Genuit, Cochair
HEAD Acoustics GmbH, Eberstrasse 30a, Herzogenrath 52134, Germany

Chair's Introduction—8:35

Invited Papers

8:40

2aNS1. Relation between loudness and annoyance over time: Implications for assessing the perception of low-frequency noise. Rhona Hellman (Inst. of Hearing, Speech, & Lang. and SLPA Dept. [106A FR], Northeastern Univ., 360 Huntington Ave., Boston, MA 02115-5000, hellman@neu.edu) and Norm Broner (Vipac Engineers & Scientists Ltd., Melbourne, Australia)

The literature suggests that, in contrast to loudness, annoyance increases over time. However, in these early measurements the annoyance exposure time usually did not exceed 120 s. In the current work, this temporal interval was extended to 3600 s (1 h). Within this time frame, both loudness and annoyance of low-frequency noise spectra containing dominant tones were measured at 12 temporal intervals by the method of successive magnitude estimation. Four noise spectra that evoked the largest annoyance-to-loudness ratios in previous reports were selected as stimuli. On average, after 1 h loudness declined by 8% and annoyance declined by 12.5% at moderate sensation levels (SL). By comparison, at 27-dB SL annoyance initially exceeded loudness by 51% up to 120 s and then, declined with time more slowly than loudness. These results indicate that, despite an 11-phon decrease in loudness level, the annoyance of a low-frequency steady noise with dominant tones below 50 Hz may initially increase over time. The annoyance increase may then be followed by a gradual annoyance decrease that mirrors the loudness decrease but at a slower rate. Taken together, the obtained relations imply, in accord with other data, that annoyance is not solely loudness based.

9:00

2aNS2. Relationships of loudness, level, and time structure in pipe organ registrations and design. Wade Bray (HEAD acoustics, Inc., 6964 Kensington Rd., Brighton, MI 48116)

Time-dependent loudnesses and sound pressure levels were measured binaurally from a variety of individual stops and multi-stop registrations of several classical and theater pipe organs covering a wide range of sizes, tonal designs, pipe scalings, wind pressures, and spatial relationships. Studies were also made of modulation in individual low-frequency notes, and in notes and chords affected by the presence or absence of tremulants. Timbres of certain solo stops were assessed in terms of the relationship of specific loudness and sound pressure spectra to their importance in creating subjective timbre. In addition to loudness and sound pressure measurement, advanced analysis by a hearing model (Sottek) and a temporal/tonal pattern-measurement algorithm (''relative approach'') were employed. Significant differences in the relationship of loudness and level were found depending principally on tonal design, pipe scaling and wind pressures on one hand, and octave-related tonal centers and timbral characteristics of individual stops on the other hand. Consideration was also given to the subjective spatial soundscape resulting from different pipe organ physical layouts and from simultaneous registrations between or among physically-separated speaking divisions.

9:20

2aNS3. Spatial representation of soundscape. Mohammed Boubezari and Jos-Luis Bento Coelho (CAPS, Intituto Superior Tecnico, 1049-001 Lisboa, Portugal, m.boubezari@ist.utl.pt)

For the last 30 years the concept of soundscape has been largely adopted in many scientific disciplines and by the urban experts for the benefit of a better comprehension and management of the sound environment. However, the spatial representation of the soundscape as a simple tool for the description, management or composition of sound environment is always needed. In this article a method is presented for the spatial sound representation with differentiated sources. The first results are shown. This method gives an account of the soundscape as close as possible to the way it can be perceived by the listener in each location. This method generates qualitative sound maps in a reduced urban scale, based on *in situ* measurements and on the implication of the measuring subject perception. The maps are sufficient enough to isolate many sound sources of the overall sound field. In this manner, sound quality refers to the sound attribute of a perceived object. It is neither an aesthetic judgment nor traditional psychoacoustics criteria. Concrete examples of application to squares in the city of Lisbon will be shown and discussed. The limits and the prospects of such a qualitative representation will also be presented and discussed.

9:40

2aNS4. Listening to a town: The urban soundscape as an image of the city. Catherine Semidor (GRECO-Bx, EAPBx, Domaine de Raba, F-33400 Talence, France, catherine.semidor@bordeaux.archi.fr)

City dwellers often have a visual representation of their town. They are able to identify some areas from pictures. But, urban sounds are also a characteristic part of the city's identity. A previous paper [C. Semidor, Proceedings 20th PLEA, Santiago, Chile (2003)] shows the possibility to distinguish the soundscape of a ''car-free'' Sunday from the soundscape of a Sunday with cars. The proposed method based on the so-called soundwalk could be very illustrative for this suggestion. These soundwalks are sound recordings made with a ''dummy'' head and a DAT recorder, on walks following a trajectory with different urban forms and, in this case study, in different towns where a tramway is running. The acoustical images (time versus frequency graphs) of the sound signals could be a representation of these soundscapes. This type of information could help urban planners and other town designers to improve the acoustic comfort of the cities. The objective of this approach is to valorize what is pleasant and agreeable in the urban sound environment.

10:00–10:20 Break

10:20

2aNS5. Context sensitive road traffic noise impact mapping—Taking the neighborhood soundscape into account. Ronny Klboe (Inst. of Transport Economics, Grensesvingen 7, N-0602 Oslo, Norway), Erik Engelien, and Margrete Steinnes (Statistics Norway, N-2225 Kongsvinger, Norway)

Road traffic noise exposure contour maps are difficult to interpret by nonexperts who are familiar with neither the road traffic noise exposure measures nor their associated impacts. An alternative is to map impacts such as annoyance. However, in urban areas the noise impacts are multi-factorially determined and context sensitive. In particular people become more annoyed by a given noise level at the most exposed facade of their dwelling when their neighborhood soundscape is even noisier. A two-tiered approach makes use of contextual soundscape information in determining noise impacts and builds contiguous neighborhood sonoscapes delimiting neighborhood areas with similar noise impacts. Neighborhood sonoscape maps facilitate a more precise targeting of local noise abatement measures, and can illustrate the impacts of noise abatement measures. With appropriate classification and class labels neighborhood sonoscapes provide an environmental labeling of the expected perceived sound quality of the neighborhood for consumers, the public and planners. Neighborhood sonoscape maps may be utilized for national stratification and subsequent two-stage cluster sampling of the population. The advantage of this approach is that focused traffic counts, extended sound modeling and monitoring of noise abatement procedures, population composition, etc. can be undertaken for a limited representative set of neighborhood sonoscapes.

10:40

2aNS6. How to boost soundscapes with tax breaks—On the application of Henry George's environmental economics to noise annoyance. Cay Hehner (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany, kskanda@mach.ut.tu-berlin.de)

The improvement of the characteristic soundscape in a given metropolitan area depends largely on a varied number of parameters not the least of which are economic. The way to sustain a large city not only economically responsible behavior patterns and activities are required, but ecologically viable ones as well. And, not coincidentally, if both patterns come into balance, a reduction of noise annoyance ensues. The cohesive set of measures that would provide the framework to ensure that the necessary economic activity is not antagonistic to (acoustic) ecology is provided by the US economist Henry George who advocated a comprehensive tax on land and all natural resources to replace eventually the bulk of all taxes checking production. New data will be provided in this paper to show the results of the application of a Georgist approach of eco-taxation to the improvement of sound quality and the reduction of noise annoyance.

11:00

2aNS7. Assessing noise pollution in sensitive areas: Soundscape analysis in an alpine valley by psychoacoustic means. Peter Lercher (Univ. of Innsbruck, Sonnenburgstrasse 16, A-6020 Innsbruck, Austria, Peter.Lercher@uibk.ac.at), Klaus Genuit (HEAD acoustics GmbH, D-52120 Herzogenrath, Germany), Urs Reichart (Inst. of Tech. Acoust., D-10587 Berlin, Germany), and Dietrich Heimann (Institut fuer Physik der Atmosphaere, D-82234 Wessling, Germany)

Alpine valleys are sensitive areas due to topography, meteorology, housing, and land-use pattern, that modify noise propagation and make protection against noise pollution rather difficult. The ''amphitheater'' effect was mentioned as explanation for deviating noise-annoyance curves and health effects observed at lower sound levels. However, detailed empirical analyses are lacking. In this study a series of simultaneous, binaural sound recordings was carried out in several cross sections of the Wipp Valley along the central-European transportation route to the Brenner pass. During 6 weeks a wide variety of day- and week times was sampled with variable wind (''Foehn'') and weather conditions (dry, rain, snow). Sound recordings were paralleled by meteorological recordings near the source and on the slope. First analyses have revealed several facts. (1) The assumption of linear sound propagation to the slope is seriously in error. (2) Tonal components from gearboxes are a significant feature in the slope recordings. (3) Low-frequency modulations make the sound more intrusive on the slope—while near the source this feature is better masked. (4) Low background sound levels ($<$30 dB,A) on the slopes are in sharp contrast with incoming sound levels (52 dB,A,L_{eq} about 1200 m from the source). (5) Meteorology leads to substantial changes in measured sound levels.

11:20

2aNS8. Transfer of knowledge from sound quality measurement to noise impact evaluation. Klaus Genuit (HEAD acoustics, Ebertstr. 30a, 52134 Herzogenrath, Germany, klaus.genuit@head-acoustics.de)

It is well known that the measurement and analysis of sound quality requires a complex procedure with consideration of the physical, psychoacoustical and psychological aspects of sound. Sound quality cannot be described only by a simple value based on A-weighted sound pressure level measurements. The A-weighted sound pressure level is sufficient to predict the probabilty that the human ear could be damaged by sound but the A-weighted level is not the correct descriptor for the annoyance of a complex sound situation given by several different sound events at different and especially moving positions (soundscape). On the one side, the consideration of the spectral distribution and the temporal pattern (psychoacoustics) is requested and, on the other side, the subjective attitude with respect to the sound situation, the expectation and experience of the people (psychology) have to be included in context with the complete noise impact evaluation. This paper describes applications of the newest methods of sound quality measurements—as it is well introduced at the car manufacturers—based on artifical head recordings and signal processing comparable to the human hearing used in noisy environments like community/traffic noise.

Session 2aPA

Physical Acoustics: Materials Characterization

Ben Dzikowicz, Chair
Coastal Systems Station, 6702 West Highway 98, Panama City, Florida 32407-7001

Contributed Papers

9:30

2aPA1. Nondestructive evaluation of multilayered structures using ultrasonic guided waves generated by an electromagnetic acoustic array transducer. Michael Quarry (Lawrence Livermore Natl. Lab., 7000 East Ave., Livermore, CA 94550, quarry1@llnl.gov)

Multilayered structures with adhesively bonded plates are a common inspection problem for nondestructive evaluation. In many cases, large areas must be inspected with limited access. One powerful method for inspecting these structures is to excite ultrasonic guided wave modes that propagate over distances of many meters and into inaccessible areas that may be embedded within another structure. The utilization of guided wave modes with horizontal shear displacement for the nondestructive evaluation of adhesively bonded plates is studied here. Modes were excited in the frequency region of 200 to 800 kHz by an electromagnetic acoustic array transducer with half-inch and quarter-inch spacing. Samples were constructed for testing that included mutlilayered plates as well as multilayered structures with a transition between a single layer plate and a multilayer plate. Notches of various sizes were placed in the bottom layer to simulate cracks and test sensitivity. Experimental data shows that an electromagnetic acoustic array was able to generate a dominant horizontal shear mode to detect the notches with sufficient signal to noise for practical inspection by tuning the frequency of excitation. Notches of 10% through-wall were detected.

9:45

2aPA2. Linear and nonlinear acoustic control of accumulated fatigue damage in steel. Alexander Sutin (Stevens Inst. of Technol., Hoboken, NJ), Yulian Kin (Purdue Univ. Calumet, Hammond, IN, asutin@stevens.edu), and Paul Johnson (Los Alamos Natl. Lab., Los Alamos, NM)

We present results of linear and nonlinear acoustic testing of steel samples with different levels of fatigue damage. The steel specimens were tested under programed cyclic loading on a fatigue testing machine and accumulated different levels of fatigue damage. No visible surface-crack formations during fatigue cycling were observed. In other words, the emphasis was placed on the characterization of continued but physically invisible in service life conditions damage in different materials and structures. Both linear and nonlinear acoustic techniques were used to assess damage accumulation. (1) Impulse resonant acoustic spectroscopy (IRAS) is based on analysis of the free-sample vibration after impact excitation. It demonstrated the increasing of the resonance frequencies and Q factor with damage accumulation. (2) Nonlinear resonant acoustic spectroscopy (NRAS) is based on measurement of the resonance response for different levels of acoustic excitation. The amplitude-dependent frequency shift for damaged steel was observed to increased with damage accumulation. (3) Nonlinear wave modulation spectroscopy (NWMS) implies the modulation of ultrasound wave by lower frequency vibration. High level of the side-band components for damaged samples were observed. The comparison of different methods is given.

10:00

2aPA3. Ultrasonic attenuation in Voronoi polycrystals. Joseph A. Turner and Goutam Ghoshal (Dept. of Eng. Mech., Univ. of Nebraska-Lincoln, W317.4 Nebraska Hall, Lincoln, NE 68588-0526, jaturner@unl.edu)

Studies of elastic wave scattering in polycrystalline media are applicable to ultrasonic materials characterization and nondestructive evaluation. Ultrasonic attenuation is used to extract microstructural parameters such as grain size and grain texture. The relations between experimental data and microstructural information require ultrasonic scattering models that often oversimplify real microstructures. Here, these relations are examined in terms of numerical simulations of elastic waves through two-dimensional polycrystals. The numerical models, based on the Voronoi polycrystal, have been shown to model accurately the microstructure of polycrystalline metals and ceramics. Methods of creating Voronoi polycrystals with appropriate microstructure, including grain elongation and grain orientation, are first discussed followed by discussions of the wave propagation. The wave simulations are based on the discretization of these Voronoi cells, using finite elements, which are then integrated directly in time using ABAQUS/EXPLICIT software. The material properties of the individual Voronoi cells are chosen according to appropriate material distributions. Here, the focus is on cubic crystals that are either statistically isotropic or transversely isotropic. Attenuation results for longitudinal waves are presented and compared with scattering theories for materials of common interest. These simulations provide insight into the attenuation models relevant for ultrasonic measurements involving polycrystalline materials.

10:15–10:30 Break

10:30

2aPA4. Using eccentricity in multilayered scatterer as mechanism for tunable phononic materials. Liang-Wu Cai (Dept. of Mech. and Nuclear Eng., Kansas State Univ., 331 Rathbone Hall, Manhattan, KS 66506, cai@ksu.edu)

The concept of multiple scattering in single scatterers, originally proposed by the author, can be utilized to analyze the scattering of elastic waves by scatterers having multiple eccentric layers. Using a dual-layer circular eccentric cylindrical scatterer as an example, it is demonstrated that, in a certain frequency range, the geometric eccentricity can significantly influence the pattern of the scattered field. The geometric eccentricity of such a scatterer can be described by three parameters: the orientation angle between the line connecting the two centers and the incident direction; the ratio of radii of the two layers, and the ratio of the distance between two centers and the radius of the outer layer. It is found that, when the wavelength is comparable to the outer diameter of the scatterer, the pattern of the scattered wave field is extremely sensitive to all three geometric parameters. Furthermore, this is the frequency range that is very close to the band gap as dictated by the lattice structure. It is expected that

when such multilayered scatterers are arranged in a lattice pattern, the eccentricity can be exploited for tuning the band-gap characteristics of the resulting phononic material.

10:45

2aPA5. Active reduction of acceleration sensitivity of quartz thickness shear resonators. Mihir Patel and Yook-Kong Yong (Rutgers Univ., 623 Bowser Rd., Piscataway, NJ 08854)

The piezoelectric quartz thickness shear resonator is a very important component in stable oscillators. The resonator is usually a rectangular or circular plate with electrodes on the top and bottom surfaces. When the resonator is subjected to acceleration, the inertial effects cause stresses and strains in the plate which in turn cause the resonant frequency to shift. A method is proposed to actively minimize the inertial stresses and strains by piezoelectric stiffening using a dc electric field bias across the electrodes. Finite element models are created to calculate the effects of acceleration on resonant frequency and compare the results with experimental data. The models are then used to calculate the effects of the dc electric field bias and demostrate the reduction of acceleration sensitivity. Results are shown for the AT- and SC-cut quartz resonators.

11:00

2aPA6. Magneto-elastic interactions in Fe–Ga alloys. G. Petculescu, H. E. Bass (Natl. Ctr. for Phys. Acoust., Univ. of Mississippi, 1 Coliseum Dr., Univ., MS 38677), T. A. Lorasso (Ames Lab., Ames, IA 50011-3020), and A. E. Clark (Clark Assoc., Adelphi, MD 20783)

The elastic anisotropy of iron's most common α-structure increases significantly when gallium is substituted for a fraction of the iron. Concurrently occurring, the tremendous boost in tetragonal magnetostriction [one order of magnitude above that of pure Fe, $\frac{3}{2}\lambda_{100}(Fe_{81}Ga_{19})$ $=400$ ppm] makes Fe–Ga alloys important for magnetoelastic applications. Measurements of the elastic tensor of Fe–Ga single crystals, with or without a magnetic field and within 4–300 K, were performed using the resonant ultrasound spectroscopy technique on 100-cut rectangular parallelepipeds. In the vicinity of 24% Ga an anomaly occurs in the magnetostriction as well as in the elastic moduli and in the magnitude of their ΔE-effect, in an apparent conspiracy to maintain the magnetoelastic energy-constants ($b_{1,2}$) "normal." [Work supported by ONR.]

11:15

2aPA7. Acousto–domain interaction in ferroelectric lithium niobate. Igor V. Ostrovskii (Dept. of Phys. and Astron., Univ. of Mississippi, University, MS 38677, iostrov@phy.olemiss.edu), Mukola M. Borovoy, Oleg O. Korotchenkov, Andrij B. Nadtochii, and Roman G. Chupryna (Kiev Shevchenko Univ., Kiev 02022, Ukraine)

First observation of the reorientation of ferroelectric domains in the lithium niobate single crystals induced by ultrasound is reported. This effect is detected from the samples, which are treated by megahertz frequency-range ultrasound with above certain threshold amplitude at room temperature. Acoustic strain amplitude causing the domain reorientation is of the order of $E10-5$. The effect is directly revealed by chemically etched crystal surfaces, and independently confirmed by the acoustically induced evolution of x-ray diffraction rocking curves and changes in acousto-electric properties of the lithium niobate crystals. The physical mechanism responsible for the interaction of the domains and ultrasound can be attributed to the mechanical stress and piezoelectric field produced by a piezo-active acoustic wave. A new effect of acousto-domain interaction in ferroelectric crystals might be taken into account for a wide variety of fundamental physical phenomena involving propagation of the acoustic waves in real crystals.

TUESDAY MORNING, 25 MAY 2004 ROYAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 2aPP

Psychological and Physiological Acoustics: Perceptual Organization of Sound (Lecture/Poster Session)

Jennifer Lentz, Chair

Speech and Hearing Sciences, Indiana University, 200 South Jordan Avenue, Bloomington, Indiana 47405

Chair's Introduction—8:00

Invited Papers

8:05

2aPP1. A really complicated problem: Auditory scene analysis. William A. Yost (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626)

It has been more than a decade since Al Bregman and other authors brought the challenge of auditory scene analysis back to the attention of auditory science. While a lot of research has been done on and around this topic, an accepted theory of auditory scene analysis has not evolved. Auditory science has little, if any, information about how the nervous system solves this problem, and there have not been any major successes in developing computational methods that solve the problem for most real-world auditory scenes. I will argue that the major reason that more has not been accomplished is that auditory scene analysis is a really hard problem. If one starts with a single sound source and tries to understand how the auditory system determines this single source, the problem is already very complicated without adding other sources that occur at the same time as is the typical depiction of the auditory scene. In this paper I will illustrate some of the challenges that exist for determining the auditory scene that have not received a lot of attention, as well as some of the more discussed aspects of the challenge. [Work supported by NIDCD.]

8:25

2aPP2. Computer models of possible physiological contributions to low-level auditory scene analysis. Ray Meddis (Dept. of Psych., Essex Univ., Colchester, Essex CO3 4SQ, UK)

Auditory selective attention is a general term for a wide range of phenomena including grouping and streaming of sound sources. While we know a great deal about the circumstances in which these phenomena occur, we have little understanding of the physiological mechanisms that give rise to these effects. Because attention is sometimes under conscious control, it is tempting to conclude that attention is a high-level/cortical function and beyond our current understanding of brain physiology. However, a number of mechanisms operating at the level of the brainstem may well make an important contribution to auditory scene analysis. Because we know much more about the response of the brainstem to auditory stimulation, we can begin some speculative modeling concerning their possible involvement. Two mechanisms will be discussed in terms of their possible relevance: lateral and recurrent inhibition at the level of the cochlear nucleus. These are likely to contribute to the selection of auditory channels. A new approach to within-channel selection on the basis of pitch will also be discussed. These approaches will be illustrated using computer models of the underlying physiology and their response to stimuli used in psychophysical experiments.

8:45

2aPP3. The perceptual organization of complex sounds by birds. Micheal L. Dent (Dept. of Physiol., 1300 University Ave., Univ. of Wisconsin, Madison, WI 53706, dent@physiology.wisc.edu)

Birds have proven to be ideal models for the perceptual organization of complex sounds because, like humans, they produce, learn, and use complex acoustic signals for communication. Although conducted in laboratory settings, measures of auditory abilities in birds are usually designed to parallel the acoustic problems faced in their natural habitats, including the location of conspecifics, discrimination among potential mates, prey localization, predator avoidance, and territorial defense. As a result, there is probably more known about hearing in birds under both natural and laboratory conditions than in any other nonhuman organism. Behavioral and/or physiological experiments on complex sound perception in birds have revealed that they exhibit serial pattern perception, can discriminate frequency changes in tones embedded within tonal patterns regardless of stimulus uncertainty conditions, segregate signals into auditory streams, and exhibit comodulation masking release. In addition, binaural experiments have revealed that birds exhibit both the cocktail party effect and the precedence effect. Taken together, these results suggest that, like humans, auditory scene analysis plays a general role in auditory perception in birds and probably other animals that must parse the world into auditory objects. [Work supported by NIH DC006124.]

9:05

2aPP4. Sensorineural hearing loss and auditory perceptual organization. Joseph W. Hall III, John H. Grose, and Emily Buss (Univ. of North Carolina at Chapel Hill, 130 Mason Farm Rd., CB 7070, 1115 Bioinformatics Bldg., Chapel Hill, NC 27599-7070, jwh@med.unc.edu)

This talk will consider the implications of sensorineural hearing loss for auditory perceptual organization. In everyday environments, the listener is often faced with the difficulty of processing a target sound that intermingles acoustically with one or more extraneous sounds. Under such circumstances, several auditory processes enable the complex waveforms reaching the two ears to be interpreted in terms of putative auditory objects giving rise to the target and extraneous sounds. Such processes of perceptual organization depend upon the central analysis of cues that allow distributed spectral information to be either linked together or split apart on the basis of details related to such variables as synchrony of onset/modulation, harmonic relation, rhythm, and interaural differences. Efficient perceptual organization must depend not only upon such central auditory analyses but also upon the fidelity with which the peripheral auditory system encodes the spectral and temporal characteristics of sound. We will consider the implications of sensorineural hearing loss for perceptual organization in terms of both peripheral and central auditory processes.

9:25

2aPP5. Auditory grouping in the perception of speech and complex sounds. Chris Darwin and Marie Rivenez (Dept of Psych., Univ. of Sussex, Brighton BN1 9QG, UK, cjd@biols.susx.ac.uk)

This talk will give an overview of experimental work on auditory grouping in speech perception including the use of grouping cues in the extraction of source-specific auditory information, and the tracking of sound sources across time. Work on the perception of unattended speech sounds will be briefly reviewed and some recent experiments described demonstrating the importance of pitch differences in allowing lexical processing of speech on the unattended ear. The relationship between auditory grouping and auditory continuity will also be discussed together with recent experiments on the role of grouping in the perceptual continuity of complex sounds.

Contributed Papers

9:45

2aPP6. Binaural release from informational masking: Results from a speech identification task. Frederick J. Gallun, Christine R. Mason, and Gerald Kidd, Jr. (Hearing Res. Ctr. and Commun. Disord., Boston Univ., 635 Commonwealth Ave., Boston, MA 02215)

This study examined how binaural cues can reduce informational masking (IM) in a speech identification task. Target and masker sentences were processed into non-overlapping frequency bands, thus limiting ''energetic masking,'' and were presented over headphones. Listeners identified key words in the target sentence. In a baseline condition, target and masker were presented monotically (SmMm), producing large amounts of IM. Binaural release from IM (i.e., improved performance re. SmMm) was observed when the target was presented monotically and the masker diotically (SmM0), suggesting that the shift of masker image away from target led to the reduction in IM. Creating large interaural differences in level (ILDs) showed that for a monaural target, binaural release, though related to ILD, occurred even when the masker image should have been lateralized at the target ear. For differences in time (ITDs) up to 600 μsec, the amount of binaural release was completely independent of ITD. For a

diotic target and binaural masker, however, release only occurred for large ITDs or ILDs. These results suggest that for monaural targets and binaural maskers, IM in a speech task can be reduced through a binaural cue that is present across the entire range of biologically plausible ITDs and ILDs.

10:00

2aPP7. Using spatialized sound cues in an auditorily rich environment. Derek Brock, James A. Ballas, Janet L. Stroup, and Brian McClimens (Naval Res. Lab., Code 5513, 4555 Overlook Ave. S.W., Washington, DC 20375)

Previous Navy research has demonstrated that spatialized sound cues in an otherwise quiet setting are useful for directing attention and improving performance by 16.8% or more in the decision component of a complex dual-task. To examine whether the benefits of this technique are undermined in the presence of additional, unrelated sounds, a background recording of operations in a Navy command center and a voice communications response task [Bolia *et al.*, J. Acoust. Soc. Am. **107**, 1065–1066 (2000)] were used to simulate the conditions of an auditorily rich military environment. Without the benefit of spatialized sound cues, performance in the presence of this extraneous auditory information, as measured by decision response times, was an average of 13.6% worse than baseline performance in an earlier study. Performance improved when the cues were present by an average of 18.3%, but this improvement remained below the improvement observed in the baseline study by an average of 11.5%. It is concluded that while the two types of extraneous sound information used in this study degrade performance in the decision task, there is no interaction with the relative performance benefit provided by the use of spatialized auditory cues. [Work supported by ONR.]

10:15

2aPP8. The probabilistic relationship between auditory stimuli and their natural sources predicts pitch. David A Schwartz and Dale Purves (Ctr. for Cognit. Neurosci., Duke Univ., Box 90999, Durham, NC 27708-0999)

The phenomenology of pitch has been difficult to rationalize and remains the subject of much debate. We here examine the hypothesis that audition generates pitch percepts by relating inherently ambiguous sound stimuli to their probable sources in the human auditory environment. A database of speech sounds, the principal source of periodic sound energy for human listeners, was compiled and the dominant periodicity of each speech sound determined. A set of synthetic test stimuli were used to assess whether the major pitch phenomena described in the literature could be explained by the probabilistic relationship between the stimuli and their probable sources (i.e., speech sounds). The phenomena tested included the perception of the missing fundamental, the pitch shift of the residue, spectral dominance and the perception of pitch strength. In each case, the conditional probability distribution of speech sound periodicities accurately predicted the pitches normally heard in response to the test stimuli. We conclude that pitch depends on an auditory process that relates inevitably ambiguous sound stimuli to their probable natural sources.

Poster Papers

All posters will be on display and all contributors will be at their posters from 10:30 a.m. to 12:00 noon.

2aPP9. Effects of frequency transposition on discrimination of recycled tonal sequences. Tomoko Hashida (Grad. School of Interdisciplinary Information Studies, Univ. of Tokyo, Japan, hashida@hc.t.u-tokyo.ac.jp), Takuro Kayahara and Takao Sato (Univ. of Tokyo, Japan)

Recycled sequences have been used for examining perception of sequences which requires temporal-order judgments of components to eliminate advantages of the initial and the terminal components in sequence recognition. The purpose of this study was to examine whether listeners could discriminate recycled sequences when the center frequency was varied. Each stimulus in this study was a random sequence of ten 200-ms pure tones divided equally in log scale within 2/3 octave. The center frequency of the standard sequence was fixed at 500 Hz, while those of test sequences were varied from 62.5 to 4000 Hz. Two test sequences were prepared for each center frequency. One had the same order as the standard; the other had a different order by reversing two contiguous components. Subjects were asked to judge which one of these two test sequences had the same order as that of the standard sequence. Results indicated that (1) listeners could distinguish the order of sequences even when the center frequency was transposed, and that (2) the range of transposition within which listeners could distinguish was limited within several octaves.

2aPP10. Asymmetrical perception of dynamic changes in rising and falling sounds. Darren Kwong (Dept. of Bioengineering, Univ. of Pennsylvania, Philadelphia, PA 19104) and Fan-Gang Zeng (Univ. of California, Irvine, CA 92697)

Previous studies have demonstrated a perceptual enhancement effect for rising sounds and interpreted at an ecological level the enhancement of rising sounds as a perceptual bias towards an approaching auditory object [Neuhoff, Nature (London) **395**, 123 (1998)]. This study has not been independently replicated and both its interpretation and mechanisms are still debatable. The present study was aimed at replicating and extending the Neuhoff study by measuring just-noticeable differences in intensity for rising tone and falling sounds. Pure tones, harmonic complex tones, and white noise were used as the stimuli. Similar to Neuhoff's study, subjects had to estimate by rating on a scale the loudness changes heard corresponding to rising or falling sounds. In addition, they had to discriminate between either a rising or falling sound and a steady-state sound or discriminate between two rising or falling sounds. Both magnitude estimation and objective discrimination data were consistent with Neuhoff's results, showing that rising sounds needed 2 and 7 dB less than falling sounds in discriminating them from a steady-state sound and from a similarly dynamically changing sound, respectively. Mechanisms related to this perceptual asymmetry will be discussed.

2aPP11. A theory of three-dimensional auditory perception. Kazi Saifuddin (Dept. of Psych., Eden College, Dhaka, Bangladesh, kazi_64@yahoo.com)

A theory of auditory dimensions regarding temporal perception is proposed on the basis of the results found from a series of experiments conducted. In all experiments, relationships were investigated between the subjective judgments and the factors extracted from the autocorrelation function (ACF) of the auditory stimuli. The factors were changed by using different properties of pure-tone, complex-tone, white-noise and bandpass-noise stimuli. Experiments by paired-comparison method were conducted in the sound proof chamber except for one in a concert hall. Human subjects were asked to compare the durations of the two successive stimuli in the pair. Subjective durations were obtained in the psychometric function. Auditory stimuli were selected to use on the basis of the measured factors of ACF as parameters. Obtained results showed significant correlation between the factors of ACF and the subjective durations described well by the theory. The theory indicates loudness and pitch as two fundamental dimensions and whether the third one is the duration of the stimuli.

2aPP12. Auditory stream segregation with multi-tonal complexes in hearing-impaired listeners. Deanna S. Rogers and Jennifer J. Lentz (Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47401, deanroge@indiana.edu)

The ability to segregate sounds into different streams was investigated in normally hearing and hearing-impaired listeners. Fusion and fission boundaries were measured using 6-tone complexes with tones equally spaced in log frequency. An ABA–ABA- sequence was used in which A represents a multitone complex ranging from either 250–1000 Hz (low-frequency region) or 1000–4000 Hz (high-frequency region). B also represents a multitone complex with same log spacing as A. Multitonal complexes were 100 ms in duration with 20-ms ramps, and- represents a silent interval of 100 ms. To measure the fusion boundary, the first tone of the B stimulus was either 375 Hz (low) or 1500 Hz (high) and shifted downward in frequency with each progressive ABA triplet until the listener pressed a button indicating that a ''galloping'' rhythm was heard. When measuring the fusion boundary, the first tone of the B stimulus was 252 or 1030 Hz and shifted upward with each triplet. Listeners then pressed a button when the ''galloping rhythm ended.'' Data suggest that hearing-impaired subjects have different fission and fusion boundaries than normal-hearing listeners. These data will be discussed in terms of both peripheral and central factors.

2aPP13. Across-channel interference and grouping. Emily Buss, Joseph W. Hall III, and John H. Grose (Univ. of North Carolina at Chapel Hill, 130 Mason Farm Rd., CB7070, 1115 Bioinformatics Bldg., Chapel Hill, NC 27599, ebuss@med.unc.edu)

Thresholds were measured for detection of an increment in level of a 500-ms, 1-kHz target tone, either in quiet or in the presence of masker tones at 0.5 and 2 kHz. Maskers were either 500 ms in duration, gated synchronously with the target, or 1 s in duration, leading the target by 500 ms. In some conditions the target and/or the maskers were amplitude modulated by multiplication with a raised cosine. The level of the target to which the increment was added was fixed at 60 dB SPL; the level of each masker was selected independently from a uniform distribution ± 8 dB *re*: 60 dB. Thresholds were higher in the presence than in quiet, despite the 1-octave spectral separation. Maximum threshold elevation occurred for synchronous target/masker gating, and with all tones (target and masker) either steady or 10-Hz amplitude modulated. Thresholds in these conditions were on the order of 10 dB [in terms of $10\log(\text{delta } I/I)$]. Performance tended to improve with introduction of gating asynchrony and with 10-Hz modulation on either target or maskers. Improvements were on average approximately 5 dB, with more pronounced effects of gating. Results will be discussed in terms of spectral cues and perceptual organization.

2aPP14. Differentiating speech and nonspeech sounds via amplitude envelope cues. Robert J. Lehnhoff, Jr., Winifred Strange, and Glenis Long (Speech and Hearing Sci., City Univ. of New York—Grad. Ctr., 365 Fifth Ave., New York, NY 10016)

Recent evidence from neuroscience and behavioral speech science suggests that the temporal modulation pattern of the speech signal plays a distinctive role in speech perception. As a first step in exploring the nature of the perceptually relevant information in the temporal pattern of speech, this experiment examined whether speech versus nonspeech environmental sounds could be differentiated on the basis of their amplitude envelopes. Conversational speech was recorded from native speakers of six different languages (French, German, Hebrew, Hindi, Japanese, and Russian) along with samples of their English. Nonspeech sounds included animal vocalizations, water sounds, and other environmental sounds (e.g., thunder). The stimulus set included 30 2-s speech segments and 30 2-s nonspeech events. Frequency information was removed from all stimuli using a technique described by Dorman *et al.* [J. Acoust. Soc. Am. **102** (1997)]. Nine normal-hearing adult listeners participated in the experiment. Subjects decided whether each sound was (originally) speech or nonspeech and rated their confidence (7-point Likert scale). Overall, subjects differentiated speech from nonspeech very accurately (84% correct). Only 12 stimuli were not correctly categorized at greater than chance levels. Acoustical analysis is underway to determine what parameters of the amplitude envelope differentiate speech from nonspeech sounds.

2aPP15. The role of spectral asynchrony in the identification of spectrally smeared environmental sounds. Valeriy Shafiro (Dept. of Commun. Disord. and Sci., Rush Univ. Medical Ctr., 1015 Armour Academic Ctr., 1653 W. Congress Pkwy., Chicago, IL 60612, valeriy_shafiro@rush.edu) and Yana Gilichinskaya (CUNY Grad. Ctr., New York, NY 10016)

This study examined possible contributions of spectral asynchrony caused by the differences in filter group delays across channels to the previously reported decline in identification accuracy of several spectrally smeared environmental sounds [V. Shafiro, J. J. Jenkins, and W. Strange, J. Acoust. Soc. Am, **113**(4), 2326–2327 (2003)]. As in the previous experiment, stimuli were obtained by processing environmental sounds through a noise-based vocoder with a varying number of frequency channels. In the present experiment, however, the group delays were held constant across channels, while the overall frequency response of the filters closely resembled the earlier one. Listeners were instructed to identify the most likely source of each sound by selecting it from 60 possible response options. Results indicated that identification accuracy improved for some environmental sounds, while it worsened for other sounds. Overall, there was little improvement in identification accuracy across channels. These findings are interpreted as reflecting the limitations in signal-processing parameters and undesirable processing artifacts. Preliminary results of a different signal-processing method that maintains a constant group delay and uses linear frequency intervals for individual channels indicate a greater improvement in identification accuracy.

2aPP16. Auditory distance perception in fixed and varying simulated acoustic environments. Matthew Schoolmaster (Hearing Res. Ctr. and Dept. of Cognit. and Neural Systems, Boston Univ., 677 Beacon St., Boston, MA 02215, msch@bu.edu), Norbert Kopčo (Boston Univ., kopco@bu.edu), and Barbara G. Shinn-Cunningham (Boston Univ., Boston, MA 02215)

Listeners must calibrate to the acoustic environment in order to judge source distance using reverberation. Results of a previous study [Schoolmaster, Kopčo, and Shinn-Cunningham, J. Acoust. Soc. Am. **113**, 2285 (2003)] showed that distance perception is more accurate when the simulated acoustic environment is consistent rather than randomly chosen from trial to trial; however, the environments were very different (anechoic versus a large classroom). The study was replicated here using two rooms that were much more similar. Each subject completed two series of trials. In the fixed-room series, the room (large or small) was fixed within a trial block. In the mixed-room series, the room was randomly chosen from trial to trial. Half of the subjects performed the mixed and half performed the fixed series first. Differences between subject groups were smaller in the current study than when anechoic and reverberant trials were intermingled; performance in the two subject groups was similar for (1) fixed trials and (2) trials simulating the large room. However, there was an interaction between subject group and room. Results suggest that listeners calibrate distance judgments based on past experience, but that such calibration partially generalizes to similar environments [Work supported by AFOSR and the NRC.]

2aPP17. Stage analysis of auditory event perception: Posture in human gait. Richard E. Pastore, Matthew J. Solomon, Jeremy R. Gaston, Jesse D. Flint, and Melody S. Berens (Dept. of Psych., Binghamton Univ., Binghamton, NY 13902-6000, pastore@Binghamton.edu)

One ultimate goal of auditory research is to understand how we perceive natural events. In our conceptualization, there are three component stages to auditory event perception: the source event, the sounds produced, and perception of the event. The current study carefully examines each of these three components, and the inter-relationships among their properties. The auditory event class for the current study is human gait in a fixed environment (e.g., shoes, walking surface, room), with posture (upright versus stooped) manipulated within individual walkers. We analyze the source events in terms of both the anthropometrical and dynamic biomechanical properties of the individual walker, and sound in terms of temporal and spectral properties of sole and heel collisions with the floor. Finally, perception is evaluated in terms of accuracy in identifying source properties (e.g., posture). In addition to statistical analyses of source–sound, sound–perception, and source–perception relationships, the source–sound link is investigated using biomechanical modeling. These analyses provide a comprehensive picture of the relationships among our three conceptual stages as well as identifying what is, and is not, relevant to perception of this specific source event. [Research supported by NSF.]

2aPP18. Temporal streaming based on fine structure cues. Nicolas Grimault (UMR CNRS 5020 Universit Lyon, 1 50 av. T Garnier, 69366, Lyon Cedex 07, France, nicolas.grimault@olfac.univ-lyon1.fr)

Alternating sequence of two *A* and *B* bursts of sound tends to be perceptually organized in different auditory streams when introducing either a spectral or a temporal difference between *A* and *B*. In general, the channeling theory of streaming predicts that any salient difference between the excitation patterns evoked by *A* and *B* sounds would lead to a segregated percept. Some previous studies have, however, evidenced that a sequence of sounds with similar spectral properties but with different temporal properties can be heard as segregated. In particular, temporal cues are probably responsible for the segregated percept when hearing a sequence of bursts of noises that are amplitude-modulated at different rates. All previous streaming experiments involved stimuli that had frequency components below 5000 Hz. As a consequence, the individual contribution of temporal fine structure cues and envelope cues could not be dissociated and remains largely undetermined. The current experiment is dedicated to test further the relative importance of both envelope and fine structure cues to segregate sequences of unresolved complex tones with different fundamental frequencies. Two high-frequency regions above and below 5000 Hz and several phase relationships leading to several temporal peak factors have been used in a subjective streaming task.

2aPP19. Mechanisms underlying the auditory continuity illusion. Daniel Pressnitzer, Julien Tardieu (CNRS UMR 9912, Ircam-CNRS, 1 place Stravinsky, F-75004 Paris, France), Richard Ragot, and Sylvain Baillet (CHU Pitié Salpétrière, F-75013 Paris, France)

This study investigates the auditory continuity illusion, combining psychophysics and magnetoencephalography. Stimuli consisted of amplitude-modulated (AM) noise interrupted by bursts of louder, unmodulated noise. Subjective judgments confirmed that the AM was perceived as continuous, a case of illusory continuity. Psychophysical measurements showed that the illusory modulation had little effect on the detection of a physical modulation, i.e., the illusory modulation produced no modulation masking. Duration discrimination thresholds for the AM noise segments, however, were elevated by the illusion. A whole-head magnetoencephalographic system was used to record brain activity when listeners attended passively to the stimuli. The AM noise produced a modulated magnetic activity, the auditory steady-state response. The illusory modulation did not produce such a response, instead, a possible neural correlate of the illusion was found in transient evoked responses. When the AM was interrupted by silence, oscillatory activity in the gamma-band range as well as slow evoked potentials were observed at each AM onset. In the case of the illusion, these neural responses were largely reduced. Both sets of results are inconsistent with a restoration of the modulation in the case of illusory continuity. Rather, they point to a role for onset-detection mechanisms in auditory scene analysis.

TUESDAY MORNING, 25 MAY 2004 LIBERTY 1/2, 9:00 A.M. TO 12:00 NOON

Session 2aSA

Structural Acoustics and Vibration: Structural Vibration

Dean Capone, Chair

Flow and Structural Acoustics, Applied Research Laboratory, Pennsylvania State University, P.O. Box 30, State College, Pennsylvania 16804

Contributed Papers

9:00

2aSA1. Distributed versus concentrated load distribution for a shell subject to a rotating constant load. Mauro Pierucci (Dept. of Aerosp. Eng., San Diego State Univ., 5500 Campanile Dr., San Diego, CA 92182-1308, mpierucci@engineering.sdsu.edu) and Scott Rosen (San Diego State Univ., San Diego, CA 92182)

When a beam or a shell is forced by a concentrated load at a discrete frequency, the structure response is composed of the sum of its natural modes. The concentrated load is always approximated by a point load as represented by the Dirac delta function. If the point load is replaced by a constant magnitude-distributed load with the same integrated value, the importance of the different modes is a function of the width of the loading function. Results will be presented for a beam and a shell. The beam is driven by a stationary load at a constant frequency. The shell is forced by a time-invariant load rotating around the shell at a frequency.

9:15

2aSA2. Dynamic response of *in-vacuo* elliptic cylindrical shells. Jeffrey E. Boisvert (Naval Undersea Warfare Ctr., Newport, RI 02841) and Sabih I. Hayek (Penn State Univ., University Park, PA 16802, boisvertje@npt.nuwc.navy.mil)

The equations of motion for the vibration of elliptic cylindrical shells of constant thickness were derived using a Galerkin approach. The elastic strain energy density used in this derivation has seven independent kine-

matic variables: three displacements, two thickness shear, and two thickness stretch. The resulting seven coupled algebraic equations are symmetric and positive definite. The shell has a constant thickness, h, finite length, L, and is simply supported at its ends ($z=0,L$), where z is the axial coordinate. The elliptic cross section is defined by the shape parameter, a, and the half-length of the major axis, l. The modal solutions are expanded in a doubly infinite series of comparison functions in terms of circular functions in the angular and axial coordinates. Damping is introduced into the shell via a complex Young's modulus. Numerical results for the drive and transfer mobilities due to surface force excitations were obtained for several h/l and L/l ratios, and various shape parameters, including the limiting case of a simply supported cylindrical shell ($a \sim 100$). Sample mode shapes were obtained at selected resonant frequencies. [Work supported by ONR and the Navy/ASEE Summer Faculty Program.]

9:30

2aSA3. Predicting the surface wall pressure frequency spectra of submerged cylindrical bodies. Y. F. Hwang, W. K. Bonness, and S. A. Hambric (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804-0030)

This paper addresses the modeling of structural excitations caused by low speed turbulent boundary layer flows. When the Corcos-type cross-spectrum or its corresponding wavevector-frequency spectrum is used as the forcing function in a homogeneous turbulent flow, the point frequency spectrum is the common multiplier to the factors representing either the cross-spectral terms or the wavevector terms. The capability of accurately predicting the point frequency spectrum is the premise of obtaining an accurate cross-spectrum or wavevector-frequency spectrum. Several point-frequency spectral models have been evaluated against the measured data obtained from buoyancy propelled cylindrical bodies. It was found that predictions using the most recently published point-frequency spectral models by Smolyakov [Acoust. Phys. **46**(3), 342–347; translated from Akusti. Zh. **46**(3), 410–407 (2000)] and Goody [AIAA-2002-2565, 8th AIAA/CEAS Aeroacoustics Conference and Exhibit, Breckenridge, CO, (2002)] provide a reasonably good agreement with the measured data. [Work supported by ONR.]

9:45

2aSA4. A theory for thermoelastic damping in MEMS and NEMS flexural oscillators. Andrew Norris (Mech. and Aerosp. Eng., Rutgers Univ., Piscataway, NJ 08854-8058, norris@rutgers.edu)

Intrinsic damping is a major roadblock in the drive to smaller high frequency RF oscillators, from the current mm size to MEMS and ultimately NEMS. Recent measurements on silicon MEMS platelike oscillators show that thermoelastic (TE) damping is the limiting damping mechanism. This talk will describe how the classical theory of Zener for TE damping needs to be revised for small thickness NEMS oscillators. A critical plate thickness is identified, separating thermally thick and thin regimes. In thermally thick oscillators the thermal diffusion is across the thickness, whereas in-plane thermal diffusion can dominate in thermally thin plates. Using the Kirchhoff assumption for the elastic deformation, it is shown that the local thermal relaxation loss depends upon the local state of vibrating flexure, specifically, the principal curvatures at a given point on the plate. The thermal loss is zero at points where the principal curvatures are equal and opposite, i.e., saddle shaped deformation. Effective plate equations are derived that incorporate TE loss as a damping term. A new form of the effective damping is obtained for both thermally thick and thin plates, generalizing the single relaxation Zener model. Applications to MEMS and NEMS oscillators will be discussed.

10:00

2aSA5. Coupled dynamic systems and Le Chatelier's principle in noise control. G. Maidanik and K. J. Becker (Code 7030, Carderock Div., Naval Surface Warfare Ctr., 9500 MacArthur Blvd., West Bethesda, MD 20817)

Investigation of coupling an externally driven dynamic system—a master dynamic system—to a passive one—an adjunct dynamic system—reveals that the response of the adjunct dynamic system affects the precoupled response of the master dynamic system. The responses, in the two dynamic systems when coupled, are estimated by the stored energies (E_s) and (E_0), respectively. Since the adjunct dynamic system, prior to coupling, was with zero (0) stored energy, $E_s^0=0$, the precoupled stored energy (E_0^0) in the master dynamic system is expected to be reduced to (E_0) when coupling is instituted; i.e., one expects $E_0<E_0^0$. In this case a beneficial noise control of the master dynamic system would result from the coupling. It is argued that the change in the disposition of the stored energies as just described may not be the only change. The coupling may influence the external input power into the master dynamic system which may interfere with the expected noise control. Indeed, the coupling may influence the external input power such that the expected beneficial noise control may not materialize. Examples of these kinds of noise control reversals are cited.

10:15–10:30 Break

10:30

2aSA6. On the ensemble variance of the energy of random plates loaded by point forces and distributed pressure fields. Vincent Cotoni, Bryce Gardner (ESI R&D, 12555 High Bluff Dr., Ste. 250, San Diego, CA 92130, vcotoni@vasci.com), and Robin Langley (Univ. of Cambridge, Cambridge CB2 1PZ, UK)

Statistical energy analysis (SEA) is used to predict the dynamic behavior of structural-acoustic systems at high frequencies. SEA predicts the mean vibrational energy over an ensemble of nominally equivalent systems. For example, a collection of individual cars coming off an assemble line would make an ensemble whose mean response is predicted by SEA. Recently a prediction of the variance of the energy in the SEA ensemble has been developed. At high frequencies where small changes in the system cause shifts in natural frequency that are larger than the average spacing of the natural frequencies, the variance becomes independent of the detailed nature of the differences between the members of the ensemble. This variance of energy can then be predicted with information available in an SEA model and with some few additional parameters describing the nature of the excitations and couplings. This new formulation is applied to an ensemble of random plates loaded by point forces and distributed pressure fields. An accurate prediction of the observed variance is demonstrated.

10:45

2aSA7. The influence of stiffened plate modifications on the energy balance. Marek Iwaniec and Ryszard Panuszka (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Eng. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland)

The change in plate stiffness introduced by a system of stiffeners slightly changes the whole system mass and results in a significant change in natural frequencies, distribution and values of vibrational energy. Stiffeners can be treated as an additional subsystem that not only changes the plate stiffness but also accumulates and dissipates a part of vibrational energy. The energy balance analysis of a stiffened plate was carried out for a plate treated as an assembly of a homogeneous plate and a system of beams. On the basis of modal participation factors the energies dissipated in plates and system of stiffeners were determined. The natural frequencies and modal participation factors were computed as a solution of an unconstrained optimization problem by the use of the Nelder Mead optimization algorithm and compared to a twice faster genetic algorithm. Theoretical computations of a homogeneous plate and stiffeners energies were verified

by energy distribution carried out on the basis of frequency response functions of FEM models. The influence of geometrical and material properties on natural frequencies and energy dissipated by plate and stiffeners was determined.

11:00

2aSA8. Dynamics of transversely isotropic cylinders. Subrata K. Bhattacharyya and Chiruvai P. Vendhan (Ocean Eng. Dept., I.I.T. Madras, Chennai, Tamilnadu 600036, India, vendhan@iitm.ac.in)

Studies on wave propagation in cylinders have a long history [M. J. Berliner and R. Solecki, J. Acoust. Soc. Am. **99**, 1841–1853 (1996)]. Onoe *et al.* [J. Appl. Mech. **29**, 729–734 (1962)] studied the attenuating and propagating modes of an isotropic cylindrical rod. Mirsky [J. Acoust. Soc. Am. **37**, 106–1021 (1965)] presented the solution for harmonic wave propagation in a transversely isotropic (TI) cylinder in vacuo. Sinha *et al.* [J. Acoust. Soc. Am. **91**, 1132–1143 (1992)] studied isotropic cylinders in contact with fluids employing the complex wave number solution. Bhattacharyya [Ph.D. thesis, Dept. of Ocean Eng., I.I.T. Madras (1989)] generalized Mirsky's work to study a TI cylinder in contact with fluids. Berliner and Solecki (cited above) also presented a similar solution and applied it to strongly TI cylinders. The present article is concerned with the vibration of a TI cylinder under prescribed harmonic end loading. The problem may be viewed as the dynamic analog of the Saint-Venant problem. The forced vibration solution is facilitated by an orthogonality relation among the radial eigenfunctions corresponding to the real and complex roots of the dispersion relation [W. B. Fraser, J. Sound Vib. **43**, 568–571 (1975)]. This approach can be used for the rigorous solution of finite cylinders.

11:15

2aSA9. High-frequency vibrations of a piezomagnetic plate. Gulay Altay (Dept. of Civil Eng., Faculty of Eng., Bogazici Univ., Bebek, 34342 Istanbul, Turkey, askarg@boun.edu.tr) and M. Cengiz Dokmeci (Istanbul Tech. Univ., Gummussuyu, 34430 Istanbul, Turkey)

This paper deals with a consistent derivation of the hierarchic system of two-dimensional approximate equations of a piezomagnetic plate at high frequency where the wavelength is of the order of magnitude or smaller than the plate thickness. To begin with, a generalized variational principle is reported for piezomagnetism, and the field variables are represented by the power series expansions in the thickness coordinate. Next, with the aid of the variational principle together with the series expansions, the system of plate equations is derived in invariant, differential and fully variational forms. The system of equations that may be readily expressed in a particular coordinate system most appropriate to the plate geometry is capable of predicting all the types of vibrations at both low and high frequency. Also, the uniqueness is investigated in solutions of the system of plate equations, and the conditions are enumerated for the uniqueness. Further, certain cases involving special geometry, material properties and types of vibrations are indicated. The resulting equations agree with and generalize some of earlier plate equations [cf. the authors, Int. J. Solids Struct. **40**, 4699–4706 (2003)]. [Work supported by TUBA.]

11:30

2aSA10. Free flexural vibrations of super elliptic plates. Murat Altekin (Faculty of Civil Eng., Yildiz Tech. Univ., Besiktas, 34330 Istanbul, Turkey, altekin@yildiz.edu.tr), Gulay Altay (Bogazici Univ., 34342 Istanbul, Turkey), and M. Cengiz Dokmeci (Istanbul Tech. Univ., 34430 Istanbul, Turkey)

This paper deals with a direct problem of the free flexural vibrations of a plate with rounded corners on the basis of the theory of thin plates. The plate of uniform thickness is made of linear elastic, isotropic materials, and its periphery is given by a super elliptic function with a power. The super elliptic power defines the shape of the plate ranging from an ellipse to a rectangle and indicates the degree of roundness. The method of solution is based on the method of separation of variables together with certain methods of weighted residuals (i.e., the method of moments, and the Galerkin and least squares methods). The shape functions are chosen in the form of double series polynomials, and they satisfy either a simply supported or a clamped boundary condition. Some numerical results are reported for the vibration frequencies in the case of both symmetric and antisymmetric mode shapes of the plate. The results are compared with the existing literature, and the convergence of solutions is discussed. [Work supported in part by TUBA.]

11:45

2aSA11. Hamilton's principle applied to piezomagnetism and related variational principles. M. Cengiz Dokmeci (Istanbul Tech. Univ., P.K.9, Gumussuyu, 34430 Istanbul, Turkey, cengiz.dokmeci@itu.edu.tr) and Gulay Altay (Bogazici Univ., Bebek, 34342 Istanbul, Turkey)

In piezomagnetism, the fundamental equations have been developed in differential form [e.g., V. I. Alshits and A. N. Darinskii, Wave Motion **15**, 265–283 (1992)]. Alternatively, they may be expressed in variational form with its well-known features; this is the topic of this paper. First, the magnetic vector, that is, the gradient of the magnetic potential, is introduced [cf. the authors, Int. J. Solids Struct. **40**, 4699–4706 (2003)]. Second, the sufficient conditions based on the energy argument are enumerated for a unique solution in the fundamental equations. Third, Hamilton's principle is stated and a three-field variational principle is obtained. The principle yields only the divergence equations and some natural boundary conditions, and it has the remaining fundamental equations as its constraint conditions. The conditions are generally undesirable in computation, and they are accordingly removed through an involutory transformation [e.g., the authors, Int. J. Eng. Sci. **40**, 457−489 (2002)]. Thus, a unified variational principle operating on all the field variables is derived in piezomagnetism. The principle is shown, as special cases, to recover some of earlier ones. [Work supported by TUBA.]

Session 2aSC

Speech Communication: Forty Years of VOT (Voice Onset Time) (Lecture/Poster Session)

Anders Lofqvist, Cochair
Department of Logopedics and Phoniatrics, Lasarettet, Lund S-221 85, Sweden

Laura Koenig, Cochair
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

Invited Papers

8:00

2aSC1. Lisker and Abramson: Teaching researchers. Lawrence J. Raphael (Dept. of Commun. Sci. and Disord., Adelphi Univ., Garden City, NY 11530, raphael@adelphi.edu)

The ''discovery'' of voice onset time (VOT) has, perhaps, stimulated more research projects than any other comparable measure. When the fathers of VOT, Leigh Lisker and Arthur Abramson, published their paper in 1964, other researchers were quick to recognize the implications of their findings and of their methodology. Those in the neighborhood of the creation in the mid-1960s were carefully trained in the proper procedures for making VOT measurements and were taught when such measures were, or were not, appropriate. In this regard, the dynamic duo was simply sustaining the tradition of using research as the basis of teaching, a tradition which both Arthur and Leigh have continued to the present. For those of us who have been fortunate enough to associate with them, their legacies as teachers stand as high in their list of accomplishments as anything else. Any academic researcher would do well to emulate their effective and often unique teaching methods, some of which will be described in this paper.

8:20

2aSC2. Producing VOT contrasts. Anders Lofqvist (Haskins Labs., 270 Crown St., New Haven, CT 06511, lofquist@haskins.yale.edu)

The development of voice onset time (VOT) as an acoustic index for studying and classifying stop consonants also prompted a large number of studies examining laryngeal activity and interarticulator timing related to VOT. A collaboration between the Research Institute of Logopedics and Phoniatrics at the University of Tokyo and Haskins Laboratories resulted in a long line of studies using electromyographic and other techniques that provided much of the empirical foundations for what we know about laryngeal function in speech, in particular the production of voiced and voiceless consonants. This presentation will review the articulatory control of VOT differences. To make a consonant voiceless, a speaker uses a combination of glottal abduction and vocal fold tensing. The distinction between voiceless stops with long and short VOT is basically due to a difference in the timing between the glottal abduction gesture and the oral closing and opening gestures. Variations in the size of the glottal gesture also occur. More generally, variations in interarticulator timing between glottal and oral movements are used to produce the different stop categories that occur in the languages of the world. [Work supported by NIH.]

8:40

2aSC3. VOT and the perception of voicing. Robert E. Remez (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027, remez@columbia.edu)

In explaining the ability to distinguish phonemes, linguists have described the dimension of voicing. Acoustic analyses have identified many correlates of the voicing contrast in initial, medial, and final consonants within syllables, and these in turn have motivated studies of the perceptual resolution of voicing. The framing conceptualization articulated by Lisker and Abramson 40 years ago in physiological, phonetic, and perceptual studies has been widely influential, and research on voicing now adopts their perspective without reservation. Their original survey included languages with two voicing categories (Dutch, Puerto Rican Spanish, Hungarian, Tamil, Cantonese, English), three voicing categories (Eastern Armenian, Thai, Korean), and four voicing categories (Hindi, Marathi). Perceptual studies inspired by this work have also ranged widely, including tests with different languages and with listeners of several species. The profound value of the analyses of Lisker and Abramson is evident in the empirical traction provided by the concept of VOT in research on the every important perceptual question about speech and language in our era. Some of these classic perceptual investigations will be reviewed. [Research supported by NIH (DC00308).]

9:00

2aSC4. VOT in cross language and comparative phonetics. Ian Maddieson (Dept. of Linguist., Univ. of California, 1203 Dwinelle Hall #2650, Berkeley, CA 94720-2650, ianm@socrates.Berkeley.edu)

Since the earliest introduction of the concept of voice onset time to describe laryngeal timing in stop consonants (Lisker and Abramson, 1964), this measure has been used as a tool for discussing phonetic similarities and differences among languages. The distinction between broad categories of leading, aligned, or lagging voice onset in relation to oral release is useful across many languages. There are also more subtle effects on VOT duration, such as some concerning major consonant places of articulation and

following vowel contexts which have proved to be repeated across languages. Combining these regularities allows relatively precise prediction of VOT in timing models. The VOT concept has also been extended to discussion of certain less-common classes of segments such as ejective stops, as well as to fricatives and affricates. The latter case raises interesting questions of distinguishing between frication noise and aspiration noise, and of what to take as the consonant's release. How best to characterize laryngeal timing around the release of segments not followed by a vowel or voiced sonorant will also be considered.

9:20

2aSC5. VOT in speech-disordered individuals: History, theory, data, reminiscence. Gary Weismer (Dept. of Communicative Disord. and Waisman Ctr., Univ. of Wisconsin—Madison, 1975 Willow Dr., Madison, WI 53706)

Forty years ago Lisker and Abramson published their landmark paper on VOT; the speech-research world has never been the same. The concept of VOT as a measure relevant to phonology, speech physiology, and speech perception made it a prime choice for scientists who saw an opportunity to exploit the techniques and analytic frameworks of ''speech science'' in the study of speech disorders. Modifications of VOT in speech disorders have been used to draw specific inferences concerning phonological representations, glottal–supraglottal timing, and speech intelligibility. This presentation will provide a review of work on VOT in speech disorders, including (among others) stuttering, hearing impairment, and neurogenic disorders. An attempt will be made to collect published data in summary graphic form, and to discuss their implications. Emphasis will be placed on how VOT has been used to inform theories of disordered speech production. I will close with some personal comments about the influence (unbeknowest to them) these two outstanding scientists had on me in the 1970s, when under the spell of their work I first became aware that the world of speech research did not start and end with moving parts.

9:40

2aSC6. VOT and hearing impairment. Harlan Lane and Joseph Perkell (Res. Lab. of Electron., MIT, Rm. 36-511, 50 Vassar St., Cambridge, MA 02139, lane@speech.mit.edu)

When deafened adults recover some hearing after receiving a cochlear implant, numerous changes in their speech occur at both phonemic and suprasegmental levels. If a change toward normative values is observed for some phonemic parameter, it may be attributed to the restored hearing; however, it may be a by-product of a suprasegmental change. Consistent with results reported for speakers with normal hearing, Lane *et al.* [J. Acoust. Soc. Am. **98**, 3096–3106 (1995)] observed in implant users that VOT varies approximately linearly with syllable duration. Therefore, in comparing pre- and postimplant measures of VOT in five speakers, each token's VOT was adjusted for the change in syllable duration of that token relative to the mean syllable duration in a baseline session (called VOTc). Preimplant, the deaf speakers characteristically uttered plosives with abnormally short VOTc. With some hearing restored, four of the five lengthened VOTc. Changes in voiced plosives' VOTc with restored hearing were correlated with changes in SPL. Some of the reliable VOTc increases that were not correlated with SPL may have been caused by auditory validation of an internal model for phoneme production. Recent studies of VOT in hearing-impaired speakers will be reviewed in this light. [Work supported by NIDCD, NIH.]

10:00–10:15 Break

Poster Papers

All posters will be on display and all contributors will be at their posters from 10:15 a.m. to 12:00 noon.

2aSC7. A comparison of English versus Spanish voicing perception using natural speech. Joan M. Sinnott and Jazmin Camchong (Psych. Dept., Univ. of South Alabama, Mobile, AL 36688)

English versus Spanish voicing perception over the past 30 years has been almost exclusively studied using synthetic VOT continua, and there has been very little research using natural VOT stimuli. This study used a balanced symmetrical design to explore the effects of training English and Spanish listeners to categorize natural tokens of English versus Spanish /b-p/ using four different vowels /i,e,a,u/. Extensive training with feedback was conducted over several months, and percent correct categorization and reaction time were analyzed. Results showed that each language group consistently exhibited enhanced performance for native speech, and this difference persisted with training. For example, reaction times leveled off at approximately 50 ms faster for native versus non-native speech. It was concluded that, while lab training can improve the ability to perceive a non-native voicing contrast, it does not result in native-like perception. Some preliminary results from monkeys using the same stimuli and procedure indicate that, unlike human adults, monkeys are more like human infants and find English and Spanish voicing contrasts equally salient. [Research supported by NIH.]

2aSC8. Voice onset time and global foreign accent in mature Japanese–English bilinguals. Timothy J. Riney (Grad. School Div. of Education, Intl. Christian Univ., Tokyo, Japan)

Experiment 1 investigates to what degree the phonetic category established in childhood for an L1 sound may evolve gradually [Flege, Speech Learning Model (1995)]. VOT is assessed in two groups of five bilingual (English and Japanese) speakers, one L1 English and the other L1 Japanese, ages 35–60 with little or no L2 exposure before age 15, but working in the L2 regularly for 15–35 years in adulthood. Both groups read lists of English and Japanese words containing word initial /p/, /t/, and /k/; 880 tokens (10 speakers $\times$ 2 languages $\times$ 44 tokens per language) were examined. The findings for both groups were that (a) the L1 VOT values did not change much over time and (b) the L2 VOT values were halfway between the L1 and target language values. Experiment 2 investigates the relationship between individual VOT and accent. The ten speakers were asked to read five Japanese sentences, now digitized. Ten additional L1 Japanese participants will rate these sentences for accent in March 2004. The prediction is that the American speakers who have the most Japanese accents will also have VOT values that are closest to L1 Japanese values.

2aSC9. The acquisition of voice onset time by English-learning infants. Joanna H. Lowenstein and Susan Nittrouer (Ctr. for Persons with Disabilities, Utah State Univ., 6800 Old Main Hill, Logan, UT 84322)

Lisker and Abramson [Word **20**, 384-422 (1964)] showed that voice onset time (VOT) for word-initial stops labeled as voiced and voiceless differ across languages, thus demonstrating that VOT is learned. Over the years, several studies have examined the developmental course of acquiring VOT for several languages, but questions remain regarding the time course of this development. Earlier studies of VOT acquisition have demonstrated that English learning infants initially produce stops with short-lag VOTs only, then go through a stage of ambiguous VOTs, finally developing two mature categories of voiced and voiceless. This study analyzed mean VOT and variability in word-initial stops produced in spontaneous samples by three infants, taped at approximately 2-month intervals between the ages of 12 months and 30 months. As predicted, these infants produced primarily voiced stops (VOT$<$40 ms) with moderate variability until approximately 21 months. Then longer lag stops began to appear and variability increased. Outliers (both voiced stops with extreme prevoicing and very long lag voiceless stops) were found. By the last recording session, all three infants showed mean voiced and voiceless VOTs approaching adult values, but still exhibited greater variability. [Work supported by NIDCD Grant No. DC-00633.]

2aSC10. Voice-onset time and buzz-onset time identification: A ROC analysis. Luis E. Lopez-Bascuas (Universidad Complutense de Madrid, Facultad de Psicologa, Campus de Somosaguas, 28223 Madrid, Spain, lelopezb@psi.ucm.es), Burton S. Rosner (Univ. of Oxford, Oxford OX1 2JF, England), and Jose E. Garcia-Albea (Universitat Rovira I Virgili, 43007 Tarragona, Spain)

Previous studies have employed signal detection theory to analyze data from speech and nonspeech experiments. Typically, signal distributions were assumed to be Gaussian. Schouten and van Hessen [J. Acoust. Soc. Am. **104**, 2980–2990 (1998)] explicitly tested this assumption for an intensity continuum and a speech continuum. They measured response distributions directly and, assuming an interval scale, concluded that the Gaussian assumption held for both continua. However, Pastore and Macmillan [J. Acoust. Soc. Am. **111**, 2432 (2002)] applied ROC analysis to Schouten and van Hessen's data, assuming only an ordinal scale. Their ROC curves suppported the Gaussian assumption for the nonspeech signals only. Previously, Lopez-Bascuas [Proc. Audit. Bas. Speech Percept., 158–161 (1997)] found evidence with a rating scale procedure that the Gaussian model was inadequate for a voice-onset time continuum but not for a noise-buzz continuum. Both continua contained ten stimuli with asynchronies ranging from -35 ms to $+55$ ms. ROC curves (double-probability plots) are now reported for each pair of adjacent stimuli on the two continua. Both speech and nonspeech ROCs often appeared nonlinear, indicating non-Gaussian signal distributions under the usual zero-variance assumption for response criteria.

2aSC11. The role of formant transitions in the perception of syllable final voicing. W. J. Warren (Dept. of Linguist., Univ. of Texas, 1 University Station B5100, Austin, TX 78712-0198) and A. E. Coren (Univ. of Texas, Austin, TX 78712-0187)

This experiment examines the role of formant transitions as a cue in syllable final voicing distinctions, specifically between low vowels and alveolar stops. Past studies have shown that when token length is a variable, listeners identify shorter tokens with voiceless codas and longer tokens with longer tokens. Ten tokens of 200 ms were synthesized from an original token and randomly played to 15 listeners. All tokens were derived from an original with steady state and formant transition period of the vowel at 100 ms each. F1 cutoff was 250 Hz and F2 cutoff was 1600 Hz. To control for token time, as transitional period was taken off in intervals of 5 ms it was then added to the steady state so that a length of 200 ms could be preserved for all tokens. Previous experiments have shown that adding 100 ms of steady state vowel onto tokens of 350 ms produces no shift in voicing distinctions, and the results here strengthen that finding. It has been shown that listeners can quickly learn to associate formant transitions with syllable final voicing distinctions when the length of the tokens is controlled.

2aSC12. Phonetic specifications in infant-directed speech. Ulla Sundberg (Dept. of Linguist., Stockholm Univ., SE-10691 Stockholm, Sweden, ulla@ling.su.se)

In contrast to the overspecification of prosodic aspects and the expansion of the vowel space typically observed in infant-directed speech (IDS), VOT measurements of voiceless stop consonants occurring in Swedish mothers' speech to their 3-month-old infants revealed surprisingly lower values in IDS than in the mothers' adult-direct speech (ADS). Since the mothers' underspecification of consonantal contrasts in early IDS, while at the same time exaggerating prosodic and vocal contrasts, may indicate an adjustment to the phonetic aspects that they perceive to be adequate to the infant's developmental level, the present study of VOT in IDS towards 12-month-olds was carried out in order to determine the age interval within which mothers may start to adopt adult-like VOT contrasts in their IDS. The VOT measurements from the IDS aimed at their 12-month-olds showed instead significantly longer durations than in ADS. It is hypothesized that this age-dependent use of consonantal specifications in IDS may reflect the mothers' adjustments to their perception of the infants' communicative capabilities, from dominating phatic function, conveyed by deeply modulated speech and exaggerated vowel contrast in early IDS, towards a focus on lexical content by the end of the first year of life.

2aSC13. Control of short lag VOT (voice-onset time) for voiced English stops. Jeff Berry (Univ. of Wisconsin—Madison, Waisman Ctr., Rm. 493, 1500 Highland Ave., Madison, WI 53705, jjberry@wisc.edu)

The American English stops /b/, /d/, and /g/ are produced with negative VOT (-150–0 ms) or short-lag VOT (0–30 ms). These two variants have been shown to occur with approximately equal frequency across speakers, although individual preferences are common. It has been hypothesized that the characteristic VOTs of short-lag stops are determined by relatively simpler articulatory control mechanisms than negative VOT variants. Specifically, if the vocal folds are approximated prior to constriction release, and intraoral pressure is vented only at the point of maximum constriction (without supraglottal expansion), the VOT will be determined by the venting process controlled by the time-varying constriction aperture. Thus, the release kinematics could account for effects of place of articulation and vowel context on short-lag VOTs. In the current work, the plausibility of this hypothesis is evaluated. The kinematics of constriction release gestures was examined from x-ray microbeam data. Descriptions of the time-varying geometries of constriction releases were input to a simple model of breath-stream dynamics and simulated VOT values were obtained. The descriptive and simulated data are consistent with the hypothesis that the characteristic durations of short-lag VOTs are determined by the venting process controlled by the supraglottal stop release gesture.

2aSC14. Perception of Spanish, Parkinsonian speech by English listeners. Michael Fraas (Dept. of Commun. Sci. and Disord., Univ. of New Hampshire, Durham, NH 03824, michael.fraas@unh.edu) and Suzanne Boyce (Univ. of Cincinnati, Cincinnati, OH 45202)

It is well established that when two languages differ in voicing patterns for obstruents, listeners are likely to misidentify obstruents in a non-native language according to the patterns of their native language. It has also been established that the normal patterns that distinguish voiced and voiceless stops in English are disrupted in speakers with Parkinson's disease. Accordingly, English-speaking clinicians not familiar with the voicing characteristics of another language, e.g. Spanish, run the risk of erroneously interpreting the utterances of clients who speak that language as natives. In a study of the voice-onset time characteristics of stops spoken

by Spanish speakers with Parkinson's disease and a matched set of control Spanish speakers (Fraas, 2002; Pan-American/Iberian Meeting on Acoustics), we determined that Spanish speakers with PD produce shorter VOTs for voiced stops and longer VOTs for voiceless stops as compared with Spanish-speaking controls. These longer VOTs were judged to be appropriate by English listeners, although they are not characteristic of normal Spanish production. These results imply that English-speaking clinicians are likely to err in not identifying voicing as a target of remediation for Spanish speakers with PD.

2aSC15. Contrastive voicing acquisition in 2-year-old children: Preliminary data. Elaine R. Hitchcock (New York Univ., New York, NY, hitchcocke@mail.montclair.edu) and Laura L. Koenig (Haskins Labs., CT 06511)

Earlier studies using voice-onset time (VOT) as the acoustic marker of contrastive voicing acquisition in English have differed widely in method and statistical procedures. Research in this area has shown three primary patterns of voicing acquisition for English stop consonants. One report indicates children demonstrate discrete voicing categories with adult-like VOT values as early as 2 years old. Other reports suggest a subperceptual distinction of the voicing contrast, followed by an exaggerated voicing contrast with evidence of more adult-like discrete categories by 3 years old. Still other reports suggest no distinction between the voicing categories until approximately 3 years. This work investigates voicing acquisition in three typically developing, English-speaking 2-year-old children. The subjects were recorded every 2 weeks for 4–6 months. Approximately 15–20 tokens were elicited for four target utterances containing initial /b p t d/. Frequency distribution, measures of central tendency, and skewness will be calculated for every recording session of each child. Discussion will focus on the development of contrastive VOT categories and their stability over time. These data will contribute to our understanding of laryngeal timing for English stop consonants in young children.

2aSC16. Physiological evidence for a temporal processing mechanism underlying voice-onset time (VOT) encoding. Mitchell Steinschneider, Yonatan I. Fishman (Dept. of Neurology, Kennedy Ctr., Rm. 322, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, steinsch@aecom.yu.edu), Igor O. Volkov, and Matthew A. HowardIII (Univ. of Iowa College of Medicine, Iowa City, IA 52242)

Despite decades of psychoacoustical research, the detailed neural mechanisms underlying VOT encoding remain obscure. Evidence collected from direct recordings in auditory cortex of human subjects undergoing surgical evaluation for medically intractable epilepsy, and from primary auditory cortex in monkeys, supports a temporal processing mechanism as a principal means by which VOT is encoded by the brain. This mechanism, as proposed by Pisoni [J. Acoust. Soc. Am. **77**, 1352–1361 (1977)], argues that the perceptual discrimination of voiced from unvoiced stop consonants is based, in part, on whether consonant release and voicing onset are perceived as occurring simultaneously or sequentially. Neural activity in auditory cortex offers physiologically plausible parallels to this perceptual scheme that can help account for the distribution of typical VOT values used by the majority of the world's languages, categorical perception of VOT, and perceptual boundary shifts that occur with changes in stop-consonant place of articulation and when nonspeech analogs of VOT are used. These responses in primary auditory cortex are poised to provide powerful inputs to later processing areas, where they can be integrated with other acoustical, visual, and language-related inputs known to modulate VOT perception. [Work supported by DC00657 and DC00120.]

2aSC17. Bilinguals' categorical perceptual shift when producing Spanish or English words: Electrophysiological correlates. Adrian Garcia-Sierra and Craig Champlin (Univ. of Texas, 1 University Station A1100, Austin, TX 78712)

In a previous study Mexican–American bilinguals identified synthetic /ga/-/ka/ syllables varying in VOT in two language conditions (Spanish and English). The perceived voicing boundary of the bilinguals varied depending on the language they had produced before the test session. In the present study event-related potentials (ERPs) were recorded from bilinguals in both language conditions. A classic oddball paradigm was used to obtain mismatch negativity (MMN) responses as an indication of categorical perception. Participants were tested in two sessions one week apart with the language conditions counterbalanced. In each test session, the electrophysiological recording was paused every 5 min in order to avoid MMN habituation. During the pause, bilinguals produced a set of Spanish or English words to keep them focused in the language of interest. The results indicated smaller MMN amplitude in the language condition where deviant and standard stimuli were more likely to be perceived as part of the same category. [Work supported by the Department of Communication Sciences and Disorders Univ. of Texas—Austin.]

2aSC18. Perception of a non-native speech contrast: Voiced and voiceless stops as perceived by Tamil speakers. Sylwia Tur (Dept. of Linguist., Univ. of Washington, Seattle, WA 98195-4340)

The effect of linguistic experience plays a significant role in how speech sounds are perceived. The findings of many studies imply that the perception of non-native contrasts depends on their status in the native language of the listener. Tamil is a language with a single voicing category. All stop consonants in Tamil are phonemically voiceless, though allophonic voicing has been observed in spoken Tamil. The present study examined how native Tamil speakers and English controls perceived voiced and voiceless bilabial, alveolar, and velar stops in English. Voice onset time (VOT) was manipulated for editing of naturally produced stimuli with increasingly longer continuum. Perceptual data was collected from 16 Tamil and 16 English speakers. Experiment 1 was an AX task in which subjects responded same or different to 162 pairs of stimuli. Experiment 2 was a forced choice ID task in which subjects identified 99 individually presented stimuli as pa, ta, ka or ba, da, ga. Experiments show statistically significant differences between Tamil and English speakers in their perception of English stop consonants. Results of the study imply that the allophonic status of voiced stops in Tamil does not aid the Tamil speakers in perceiving phonemically voiced stops in English.

Session 2aSPa

Signal Processing in Acoustics and Underwater Acoustics: Acoustic Communications—Space/Time Coding, Modulation, Processing and Propagation Effects

James C. Preisig, Chair

Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Bigelow 207, Woods Hole, Massachusetts 02543-1053

Chair's Introduction—8:00

Invited Papers

8:05

2aSPa1. Acoustic communication systems and space/time processing. Arthur B. Baggeroer (MIT, Cambridge, MA 02139)

Acoustic communication research for the last three decades has concentrated on point-to-point systems. Efforts are now evolving which exploit the spatial aspects of the underwater channel. First, there are systems with array for transmitters and/or receivers. Because the ocean randomizes the propagation, information can be received over many ''paths;'' for this beamforming and, more generally, singular value decomposition methods have been used to increase channel capacity. With multiple nodes in a distributed system, the problems of (i) time latency due to the relatively slow speed of sound and (ii) range, or multipath, spread become much more problematic; consequently, the space/time architecture of using time, frequency and space becomes very important and this differentiates underwater systems from cellular phone technology. The network structure, e.g., centralized or link connected leads to different signaling protocols and the potential for complicated interaction and interference. Nevertheless, many advances including both space/time diversity and coding are now popular research topics which will lead to more reliable and higher data rate acoustic communication systems.

8:25

2aSPa2. Optimization of transmit/receive array processing for high rate acoustic communications. Milica Stojanovic (MIT, Cambridge, MA 02139)

A broadband acoustic communication link between a user equipped with a single transmit/receive element and a station equipped with an array is considered. System optimization is conducted to obtain transmit/receive filtering techniques that provide best performance over a multipath channel under varying system constraints. Two classes of systems are considered: those that aim for complete intersymbol interference suppression and those that use equalization at the receiver side. Both systems are considered with or without the possibility to implement a perfect channel feedback. Transmit/receive filters are derived analytically, and performance of each scheme is assessed through its data detection SNR. Performance of various techniques is compared analytically, using an example of a shallow water channel, where transmission at high bit rate results in extensive delay spread.

Contributed Papers

8:45

2aSPa3. Simultaneously multiple-depth coherent communications using time reversal. Heechun Song, W. S. Hodgkiss, W. A. Kuperman, Philippe Roux, T. Akal (MPL/SIO, 9500 Gilman Dr., La Jolla, CA 92093-0238), and M. Stevenson (NATO SACLANT Undersea Res. Ctr., La Spezia 19138, Italy)

A recent time reversal experiment demonstrated that multiple foci can be projected from an array of sources to the same range but at different depths. This ''multiple input, multiple output (MIMO)'' process potentially can improve the information data rate. This paper presents the experimental results of binary phase shift keying (BPSK) communications at 3.5 kHz with a 500-Hz bandwidth where two different messages were sent simultaneously to different depths at 9-km range in 110-m-deep shallow water.

9:00

2aSPa4. Multichannel time reversal communications in a highly reverberative environment. James V. Candy, Brian Guidry, Andrew Poggio, and Claudia Kent (Univ. of California, Lawrence Livermore Natl. Lab., P.O. Box 808, L-156, Livermore, CA 94551, candy1@llnl.gov)

The development of point-to-point time-reversal (T/R) communications in a highly reverberative environment was discussed previously [Candy *et al.*, J. Acoust. Soc. Am. Suppl. 1 **114**, 2367]. This paper focuses on the extension of that effort to the multichannel case. Here, we discuss the theoretical development of a suite of multichannel TR receiver realizations, similar to the point-to-point case, using an acoustic T/R array and a set of client stations. The performance of these processors on both simulated and experimental data is discussed in detail. The experiment is provided by a stairwell between two floors of a noisy building. The stairwell is populated with obstructions (pipes, rails, wall, etc.) and a 90-deg landing—clearly a highly reverberative environment. It is shown that the

multichannel receivers perform quite well when compared to their point-to-point counterparts, and are able to reliably extract the transmitted code from the noisy measurements.

9:15

2aSPa5. Time-reversal communication through a highly reverberant medium. David H. Chambers, Claudia A. Kent, and Alan W. Meyer (Lawrence Livermore Natl. Lab., P.O. Box 808, Livermore, CA 94551)

An ultrasonic time-reversal array system is used to transmit communication signals across an aluminum slab with a large number of holes drilled through it. The hole pattern was designed to greatly attenuate the direct propagation path between each transmitter and receiver, thereby forcing the communication signal to be carried primarily by the random, scattered field. Four approaches to time-reversal communications [see Candy *et al.*, J. Acoust. Soc. Am. Suppl. 1 **114**, 2367 (2003)] were used to establish clear (negligible symbol error) communication channels through the slab. Both point-to-point and array-to-point configurations were tested. Comparisons between these approaches are shown using both simulated and experimental measurements. [Work performed under the auspices of the U. S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.]

9:30

2aSPa6. Using acoustic orthogonal signals in shallow water time-reversal applications. Thomas Folegot (Atlantide, Technopole Brest Iroise, CS 23866, 29238 Brest, France), Philippe Roux, William A. Kuperman, William S. Hodgkiss, Hee Chun Song, Tuncay Akal (Scripps Inst. of Oceanogr., La Jolla, CA), and Mark Stevenson (NATO SACLANT Undersea Res. Ctr., La Spezia 19138, Italy)

Orthogonal broadband signals (adaptive instant record, AIR Signals) are constructed that are particularly suitable to shallow water multipath propagation. They have similarities with the multi-carrier code division multiple access (CDMA) methods which have arisen from mobile telephony, and allow communication between several sources and several receivers. The period of the orthogonal signals is related to the channel time spreading in order to provide an exact and simultaneous measure of the transfer function between all sources and receivers on the whole transducer frequency band. This method which optimizes the signal-to-noise ratio is relevant to fluctuating environments and particularly to time reversal applications. The first experimental demonstration at the ultrasonic scale has shown the efficiency of these signals in a stationary environment as well as their robustness to absorption and to bottom reverberation. This presentation describes the experimental use of these signals with a time reversal mirror in the Mediterranean Sea during an April 2003 experiment.

9:45

2aSPa7. Adaptive instant record signals applied to shallow water detection. Thomas Folégot (Laboratoire Ondes et Acoustique, Université Denis Diderot, UMR CNRS 7587, ESPCI, 10 rue Vauquelin, 75231 Paris Cedex 05, France and Atlantide, Marine Sci. and Technol. Dept., Technopôle Brest-Iroise, CS23866, 29238 Brest Cedex 3, France), Julien de Rosny, Claire Prada, and Mathias Fink (Université Denis Diderot, 75005 Paris, France)

Time reversal arrays are becoming common tools whether for detection, tomography or communication. These applications require the measurement of the response from the array to one or several receivers. The most natural way to record different impulse responses between several points is to generate pulses successively from each emitting point and directly record all the impulse responses on the recording points. However, this method is very time consuming and inefficient in terms of signal-to-noise ratio. Hence, in this work, we propose an original way of sending continuous signals simultaneously from all the sources, recording all the pressure fields on the receivers and processing them in order to extract the exact impulse responses by matched filter. To this end, the signals are adapted to the environment and, more specifically, to highly dispersive media. These adaptive instant records signals (AIRS) are used experimentally to detect targets using the time reversal operator decomposition method. The quality of the 15×15 transfer functions acquired simultaneously, and therefore, the detection capability is demonstrated in shallow water in the presence of bottom absorption and reverberation. Finally, the connection of AIRS with CDMA and FDMA that are two coding techniques used in telecommunication is shown.

10:00–10:15 Break

10:15

2aSPa8. Acoustic video streaming in a waveguide. Philippe Roux, Jit Sarkar, W. A. Kuperman, W. S. Hodgkiss, Hee Chun Song, Tuncay Akal (Marine Physical Lab., Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0238, jit@mpl.ucsd.edu), and Mark Stevenson (NATO SACLANT Undersea Res. Ctr., La Spezia 19138, Italy)

The use of two vertical arrays (a source and a receiver array) yields the recording of the broadband tranfer matrix in a waveguide. Using time-reversal focusing in this geometry, we have experimentally demonstrated at sea the ability to focus on any receiver array element. Furthermore, we have also shown the possibility to achieve simultaneous multiple foci on different elements of the receiver array. Much like the old dot-matrix printers that used linear arrays of pins to create pixels on paper and thus form an image, we now investigate the use of multiple focal spots on the receiver array elements to create pixels in space-time. Vertical resolution is limited by the number of elements in the receiver array, while horizontal resolution is limited by the frequency bandwidth of the transducers. This acoustic TV opens up the possibility of transmitting full-motion video through an acoustic waveguide, which is a potentially robust form of communication as the human eye is able to recognize images even after considerable loss in quality.

10:30

2aSPa9. KauaiEx: Environmental effects on HF acoustic propagation with application to communications. Michael B. Porter, Paul Hursky, Martin Siderius (Ctr. for Ocean Res., Sci. Applications Intl. Corp., 10260 Campus Point Dr., San Diego, CA 92121), Mohsen Badiey (Univ. of Delaware), Jerald Caruthers (Univ. of Southern Mississippi), William S. Hodgkiss, Kaustubha Raghukumar (Scripps Inst. of Oceanogr.), Dan Rouseff, Warren Fox (Univ. of Washington), Christian de Moustier, Brian Calder, Barbara J. Kraft (Univ. of New Hampshire), Keyko McDonald (Space and Naval Warfare Systems Ctr., San Diego, CA), Peter Stein, James K. Lewis, and Subramaniam Rajan (Sci. Solutions, Inc.)

The Kauai Experiment (22 June–9 July 2003) was designed to study high-frequency (8–50 kHz) acoustics in a shallow-water waveguide. In contrast to much of the previous literature, emphasis was placed on multipath arising from multiple boundary interactions. Various participants were interested in different applications; however, a core theme was the role of the environment on acoustic communications. A great deal of effort was made to characterize the environment including the surface wave spectrum, 2D temperature structure along the propagation path, salinity, currents, and bottom properties. Most of these parameters were measured continuously over the 2 weeks of the experiment, providing information on the diurnal cycles. At the same time, extensive acoustic measurements were made using a variety of vertical line arrays, some of which spanned the entire water column. The acoustic measurements included channel probes to characterize the variation of the impulse response. These probes were interleaved with a variety of modulation schemes for communications including noncoherent methods such as MFSK (multifrequency shift keying), and DPSK (differential phase-shift keying), as well as coherent schemes such as QAM (quadrature amplitude modulation), OFDM (orthogonal frequency division modulation), and PPC (passive-phase conjugation) methods. Thus, the experiment provides a vast amount of information relating environment to acoustic propagation to modem performance. This talk will present an overview of key lessons learned to date.

10:45

2aSPa10. Relating ocean dynamics and sea state to time-angle variability of HF waveforms. Mohsen Badiey (Univ. of Delaware, Newark, DE 19716), Stephen E. Forsythe (Naval Underwater Warfare Ctr., Newport, RI 02841-1708), Michael B. Porter (Sci. Application Intl. Corp., San Diego, CA 92121), and The KauaiEx Group[a)]

One of the objectives of the Kauai experiment was a better understanding of the ocean dynamics effects on the propagation of high-frequency acoustic signals. Due to a unique oceanographic feature of the shallow water region near the Pacific Missile Range Facility in Kauai, a bottom mounted vertical line array containing eight elements was deployed with sufficiently small element spacing to measure the acoustic energy near the bottom. Simultaneous environmental parameters including current, temperature and salinity profiles, directional surface wave spectra, as well as the wind speed and direction above the sea surface were measured. High correlation between the environmental variability and the received acoustic signals is observed. To interpret the results broadband PE and Gaussian beam ray tracing models were utilized. Arrival time-angle statistics are correlated with the environmental variability due to ocean dynamics in this region. It is shown that variations of the sea surface dynamics exhibit different statistical effects than those occurring within the water column. [Work supported by ONR-321OA.] [a)]Paul Hursky, Martin Siderius (SAIC), Jerald Caruthers (USM), William S. Hodgkiss, Kaustubha Raghukumar (SIO), Dan Rouseff, Warren Fox (APL-UW), Christian de Moustier, Brian Calder, Barbara J. Kraft (UNH), Keyko McDonald (SPAWARSSC), Peter Stein, James K. Lewis, and Subramaniam Rajan (SSI).

11:00

2aSPa11. Impact of thermocline and seabed variability on underwater acoustic communications. Martin Siderius, Michael Porter (SAIC, Ctr. for Ocean Res., 10260 Campus Pt. Dr., San Diego, CA 92121), Finn Jensen (SACLANT Undersea Res. Ctr., 19138 La Spezia (SP), Italy), and The Kauai Group (SAIC, SPAWAR, UDel, APL-UW, MPL-UCSD, SSI, UNH, USM)

Shallow water acoustic communications experiments were conducted near Kauai in July 2003 and near Capraia and Elba Islands in October 2003. All experiments took place in approximately 100-m water depth but the oceanography and seabed types differed significantly. The Kauai site had a reflective seabed that was combined with highly variable oceanographic conditions that led to performance closely tied to source/receiver geometry. The Capraia site also has a reflective seabed but the winter conditions produced a more mixed water column with a weaker and less variable thermocline than Kauai. The Elba site had nearly the same oceanographic conditions as Capraia but the seabed is a highly lossy. In each of these experiments, signals were transmitted over many hours from fixed and moving platforms and were received at multiple ranges and depths using vertical arrays and single hydrophones. Extensive environmental measurements were made simultaneous to the acoustic transmissions (e.g., measurements of the water column temperature structure and surface wave heights). In this paper, the correlation between environmental factors and communications performance will be presented along with the predictions from modeling. The performance of both multi-frequency shift keying (MFSK) and direct sequence spread spectrum (DSSS) signals will be discussed.

11:15

2aSPa12. Comparison of two underwater acoustic communications techniques for multi-user access. Paul Hursky, T. Martin Siderius, Michael B. Porter, and The KauaiEx Group[a)] (SAIC, 10260 Campus Point Dr., San Diego, CA 92121)

Frequency hopped frequency shift keying (FHFSK) and code division multiple access (CDMA) are two different modulation techniques for multiple users to communicate with a single receiver simultaneously. In July 2003, these two techniques were tested alongside each other in a shallow water coastal environment off the coast of Kauai. A variety of instruments were used to measure the prevailing oceanography, enabling detailed modeling of the channel. The channel was acoustically probed using LFM waveforms and m-sequences as well. We will present the results of demodulating the FHFSK and CDMA waveforms and discuss modeling the channel for the purpose of predicting multi-user communications performance. [a)]Michael B. Porter, Paul Hursky, Martin Siderius (SAIC), Mohsen Badiey (UD), Jerald Caruthers (USM), William S. Hodgkiss, Kaustubha Raghukumar (SIO), Dan Rouseff, Warren Fox (APL-UW), Christian de Moustier, Brian Calder, Barbara J. Kraft (UNH), Keyko McDonald (SPAWARSSC), Peter Stein, James K. Lewis, and Subramaniam Rajan (SSI).

11:30

2aSPa13. The effect of internal waves on mid-frequency underwater acoustic communications. T. C. Yang (Naval Res. Lab., 4555 Overlook Ave., Washington, DC 20375)

Internal waves are abundant in coastal waters and are known to have a significant impact on acoustic signal propagation. The effect includes significant signal fading to loss of signal coherence depending on the acoustic signal frequencies. This paper presents the analysis results using binary phase-shifted keying signals for underwater acoustic communications during the ASCOT01 experiment. We report the measurements of temporal coherence time as a function of time and relate it to the performance of the decision feedback equalizer. The interpacket coherence was measured using the LFM signals in each packet which were separated by ~ 2 min. We see that the intensity and arrival time of the multipath arrivals vary significantly from packet to packets. The signal coherence within a packet (intrapacket) is measured using consecutive m-sequences. We find that at mid (2–5 kHz) frequencies, the acoustic environment, while randomly changing with time on the scale of minutes, presents an instantaneous deterministic environment for each packet; the coherence time is much longer than the packet length. The environmental impact can be mitigated by adequate sampling of the channel impulse response function. [Work supported by the ONR.]

11:45

2aSPa14. The effect of internal waves on low-frequency underwater acoustic communications. T. C. Yang (Naval Res. Lab., 4555 Overlook Ave., Washington, DC 20375)

Internal waves are abundant in coastal waters and are known to have a significant impact on acoustic signal propagation. The effect includes significant signal fading to loss of signal coherence depending on the acoustic signal frequencies. This paper presents the analysis results using binary phase-shifted keying signals for underwater acoustic communications (ACOMMS) during the SWARM95 experiment. As reported in the open literature, the acoustic data were strongly affected by the presence of internal waves. For example, at 200–400 Hz, the signal characteristics changes drastically at the scale of ~ 2–6 min. The most noticeable changes include the depth where the signal intensity peaks, the multipath arrival (time) pattern, and the signal amplitude and phase at a fixed depth. At these frequencies, the ACOMMS packet length for a reasonable amount of data will be long (due to the limited bandwidth) and comparable to the signal coherence time. The bit error rate of a packet varies depending on the temporal coherence within a packet (intrapacket). We report the temporal coherence measurements as a function of packet transmission time and relate that to the performance of the channel equalizer. [Work supported by the U.S. ONR.]

Session 2aSPb

Signal Processing in Acoustics and Acoustical Oceanography: Inverse Problems in Seismic Signal Processing

Max Deffenbaugh, Cochair
Exxon Mobil Upstream Research, P.O. Box 2189, Houston, Texas 77252

Alan W. Meyer, Cochair
Lawrence Livermore National Laboratories, 700 East Avenue, Livermore, California 94550

Chair's Introduction—9:00

Invited Papers

9:05

2aSPb1. Overview of inverse problems in acoustic and seismic signal processing. Leon H. Sibul and Michael J. Roan (Appl. Res. Lab., Penn State Univ., P.O. Box 30, State College, PA 16804-0030, lhs2@psu.edu)

An overview of inverse problems in acoustic and seismic signal processing is presented. The main goal of seismic inversion is to determine earth properties from seismic reflections or signals. Acoustic inversion includes ultrasonic imaging, acoustic tomography, acoustic oceanography, and propagation medium modeling. System identification, parameter estimation, image reconstruction from incomplete data, and blind deconvolution are all inverse problems. Most practical inverse problems are underdetermined and their solutions are not unique. The nonuniqueness is addressed by maximum entropy, minimum relative entropy methods, and Bayesian methods. Bayes method determines maximum *a posteriori* (MAP) probability density estimates of the model. Many inverse problems are numerically unstable. Various regularization methods are used to alleviate numerical instability. Classical least-squares inversion, singular value decomposition, and recent developments of blind deconvolution are reviewed. [Work supported by ONR.]

9:25

2aSPb2. Detection of uncertain seismic signals. D. B. Harris (Energy and Environment, Earth Sci. Div., Univ. of California, Lawrence Livermore Natl. Lab., 7000 East Ave., Livermore, CA 94550-9234, harris2@llnl.gov)

Current practice in seismic event detection is concentrated at the extremes of a spectrum of possibilities determined by the degree of knowledge available about the signals to be detected. At one end of the spectrum, information is available about only the frequency content of the signal. Correspondingly, simple energy detectors operating in fixed passbands are the detection method of choice: the seismological standard for detection of events over broad regions. At the other end of the spectrum, the signal is completely known. For such cases, correlation detectors (matched filters) are indicated. Correlation detectors are the emerging standard for detection of repetitive events from sources with very limited geographic extent (a few wavelengths), e.g., mines, compact earthquake swarms, and aftershock sequences. It is desirable to seek detectors that operate between the extremes, i.e., that have much of the sensitivity of correlation detectors, but more of the flexibility of energy detectors to detect signals with greater uncertainty. Subspace detectors offer one approach to achieving this trade-off. This talk describes an application of subspace detectors to detect events in a swarm with significant signal variability, and demonstrates the theoretical ability of these detectors to grade almost continuously between correlators and energy detectors.

9:45

2aSPb3. Characterization of seismic signals and background seismic noise. Robert Uhrhammer (Berkeley Seismological Lab., 215 McCone Hall, Univ. of California, Berkeley, Berkeley, CA 94720-4760)

A variety of signal processing methodologies have been utilized in the analysis and characterization of the seismic signals and the background noise recorded by modern seismic instrumentation at the UC Berkeley Seismological Laboratory. The capabilities of seismic instrumentation have vastly improved over the past couple of decades since the advent of microcomputer and related technologies. Modern seismic instrumentation consists primarily of force fedback broadband seismometers (to record weak ground motions) and strong motion accelerometers (to record large local shocks) coupled to high resolution digital (24-bit+integer) data loggers. We discuss the characteristics and capabilities of the various seismic sensors and the digital recording systems. The earth's background noise and the instrumental self-noise characteristics are discussed. Various weak signals (both coherent and transient) detection and analysis methodologies are also discussed along with some interesting cases of detecting and characterizing unusual natural and artificial seismic sources.

10:05–10:20 Break

10:20

2aSPb4. Approximate methods for time-reversal processing of large seismic reflection data sets. James Berryman (Lawrence Livermore Natl. Lab, P.O. Box 808, L-200, Livermore, CA 94551-9900)

Seismic reflection data sets, whether collected on land or at sea, contain very large amounts of data even by modern standards. These data are typically sampled regularly in time but irregularly in space. So, temporal Fourier transforms are straightforward, but spatial transforms can be problematic, and very often require special efforts to resample the data onto a uniform grid in space. On the other hand, time-reversal processing of such data requires only the temporal Fourier transform, which is easy for these data. Accounting for data reciprocity, a complex square transfer matrix can be constructed at each frequency, and standard methods can then be used—at least in principle—to find singular values and singular vectors (SVD) of this large matrix. Once those singular vectors having the largest singular values are known, it is relatively straightforward to reconstruct the scattering surfaces giving rise to the data matrix. The biggest leap in this data processing scheme comes from the need to find the SVD for very large complex matrices. But iterative schemes using Krylov space methods can be applied to resolve this difficulty and produce a relatively small number of singular vectors containing the most relevant information.

10:40

2aSPb5. Planarly layered diffraction tomography with accurate Green function. Sean K. Lehman (Lawrence Livermore Natl. Lab., L-154, 7000 East Ave., Livermore, CA 94550)

Diffraction tomography (DT) imaging techniques require knowledge of the background Green function. Due to its simplicity, it is standard practice to use a homogeneous medium Green function. We have developed a model of a planarly layered Green function that can be used in DT imaging of planarly layered media where the layer acoustical properties and dimensions are known. We present the theory and applications. [Work performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.]

11:00

2aSPb6. Estimating lithology and fluid parameters from seismic data. Michael E. Farrell (Exxonmobil Res. Co., P.O. Box 2189, Houston, TX 77252-2189, mike.e.farrell@exxonmobil.com)

A constrained inversion approach is presented for making quantitative estimates of lithology and fluid parameters using compressional wave reflection data (P-to-P) and converted wave reflection data (P-to-S). The method is horizon based, and requires the seismic amplitude information from multiple offsets at a particular zone of interest. Results are discussed from two specific cases. In the first example estimates of lithology are derived from only P-to-P reflection data, assuming the fluid properties are known. This turns out to be a well-posed problem with a unique solution, and is robust in the presence of additive random noise. In the second example both lithology and fluid parameters are estimated using a joint inversion of P-to-P and P-to-S seismic data. This is a much more difficult problem, and does not have a unique solution. However, by including realistic lithology and fluid constraints, good estimates of the parameters can be obtained. Furthermore, adding these constraints to the inversion of only P-to-P wave data also yields a good solution. This suggests that the constraints are more important than doing a joint inversion with the additional P-to-S data.

Contributed Papers

11:20

2aSPb7. Utilization of the interference invariant for geoacoustic parameter inversion in shallow water. Peter M. Daly and Peter N. Mikhalevsky (SAIC Ocean Systems Operation, 1710 SAIC Dr., MS 1-11-15, McLean, VA 22102, peter.m.daly@saic.com)

S. D. Chuprov's interference invariant, or β parameter, characterized broadband striation patterns as a function of range and freuencey. Puchenkina and Salin proposed using the β parameter to estimate geoacoustic bottom parameters. Further work by Baggeroer, Rousseff, and Spindel recommended characterizing β as a probability distribution instead of a discrete quantity. the authors of this presentation will illustrate use of the β distribution to perform geoacoustic parameter inversion in shallow water, using simulated annealing methods. This inherently broadband process utilizes ships of opportunity combined with *a priori* range information to provide high-SNR input signals to the inversino process. [Work sponsored by DARPA Advanced Technology Office, Contract N00024-01-C-6319.]

11:35

2aSPb8. Techniques for estimating sediment attenuation from chirp sonar data. Altan Turgut (Naval Res. Lab., Acoust. Div., Washington, DC 20375)

Recent broadband (50–500 Hz) matched-field inversions of geoacoustic parameters indicated that sediment attenuation is difficult to estimate especially at short source-receiver distances. On the other hand, geoacoustic inversions become less practical at long source-receiver distances due to the range-dependency of both ocean bottom and water column. Several inversion techniques are investigated to estimate *in-situ* sediment attenuation from normal incidence reflection data, collected by chirp sonars. Both frequency-shift and attenuation roll-off measurement techniques showed nonlinear frequency dependency of attenuation for sandy sediments at the 2-12-kHz frequency band. Estimated attenuation values are extrapolated to lower frequencies by using the Biot theory. The extrapolated results are consistent with those of the matched-field inversions. Several advantages and limitations of the frequency-shift and attenuation roll-off measurement techniques are also discussed. [Work supported by ONR.]

Session 2aUW

Underwater Acoustics: General Underwater Acoustics

Natalia Sidorovskaia, Chair
Department of Physics, University of Louisiana at Lafayette, Lafayette, Louisiana 70504

Chair's Introduction—7:55

Contributed Papers

8:00

2aUW1. New methods in the analysis of resonances. H. Uberall (Dept. of Phys., Catholic Univ. of America, Washington, DC 20064) and M. Werby (NRL Code 7181, Stennis Space Ctr., MS 39529)

From water inclusions to layered elastic shells with included fluids, all may be analyzed by a common method. We demonstrate that the background method is always possible and is because solutions of linear nonhomogeneous, differential equations are composed of the supposition of the homogeneous solution plus a particular solution. Moreover, the partial wave solutions that arise in the usual developments have both inertial and damping components. We show that one may use these facts to determine both the resonance locations and the resonance widths and we present a context for which one may give meaning to observed resonance features.

8:15

2aUW2. Acoustical imaging of underwater objects using the bistatic ramp response signals. Wei Li (Natl. Univ. of Singapore, Singapore), Gui-Rong Liu (Natl. Univ. of Singapore, Singapore), and Vijay Varadan (Penn State Univ., University Park, PA 16801, vjvesm@engr.psu.edu)

The T-matrix methods have been successfully employed for underwater acoustical imaging techniques and the inverse analysis of acoustic scattering problems. Most inverse techniques use the backscattering signals in the acoustical far field to retrieve the shape and size information of an underwater object, such as the ramp response technique. This paper addresses a modified ramp response technique, which could be used to reconstruct the 3D image of an object for the bistatic case. This technique shows that the bistatic ramp response is proportional to the profile function of an underwater object based on the small bistatic angle assumption. The numerical examples demonstrate that the bistatic ramp response technique is still valid to obtain an excellent profile function even for the bistatic case with a fairly big bistatic angle. This bistatic ramp response technique allows us to reconstruct the 3D image of an underwater object with only one receiver.

8:30

2aUW3. A theoretical method to extract target information between two vertical arrays. William Sanders and Michael Werby (NRL Code 7181, Stennis Space Ctr., MS 39529)

For a known source and suitable environmental information one may determine the presence (or absence) of a target ensonified by a guided wave initiated by the source by using vertical arrays that sandwich a region of interest. We show first that one can represent the ensonified object as a source term in the solution of an inhomogeneous Sturm–Louiville problem where the homogeneous solution corresponds to the solution in the absence of a target. This allows one to choose a bounded region about the object and evaluate the equivalent of Huygens' integral (over the boundary that includes the vertical arrays), which enables one to trace the signal back to the object. This determines both the range and depth and even the target strength. When this is extended to a pulse signal then the information leads to higher fidelity than the frequency results due to phase averaging of fluctuations. Thus we demonstrate this to be a good strategy for an experimental study that would lead to a robust detection method that may be implemented on a rapid basis.

8:45

2aUW4. The diffraction of leaky Rayleigh waves at the extremity of a fluid-loaded plate. Nico F. Declercq, A. Teklu, M. A. Breazeale (Natl. Ctr. for Physical Acoust., The Univ. of Mississippi, Oxford, MS 38677, NicoF.Declercq@UGent.be), Rudy Briers (KATHO, B-8820 Torhout, Belgium), Oswald Leroy (KU Leuven Campus Kortrijk, B-8500 Kortrijk, Belgium), and Joris Degrieck (Ghent Univ., B-9000 Ghent, Belgium)

This study reveals that leaky Rayleigh waves, when scattered at the extremity of a thick plate swamped in water, generate leaky Rayleigh waves that propagate around the corner. Furthermore, it is experimentally proved that leaky Rayleigh waves are stimulated by the borders of a Gaussian-bounded beam and not by the interior of the beam. A comparison between the scattering of leaky Rayleigh waves and the scattering of Scholte–Stoneley waves at the extremity is also outlined and shows that leaky Rayleigh waves on the vertical edge at the extremity of the plate are best stimulated by means of incident leaky Rayleigh waves and not by means of Scholte–Stoneley waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.) and by a NATO Collaborative Linkage Grant.]

9:00

2aUW5. The dependence of Lamb wave stimulation parameters on the impedance difference between upper and lower liquid. Nico F. Declercq, Filip Van den Abeele, Joris Degrieck (Soete Lab, Dept. of Mech. Constr. and Prod., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, NicoF.Declercq@UGent.be), and Oswald Leroy (IRC-KULAK, B-8500 Kortrijk)

It is known that inhomogeneous plane waves are better capable of stimulating Lamb waves in a solid plate than homogeneous plane waves. For each Lamb mode there exists an angle of incidence and an inhomogeneity parameter that maximizes its stimulation. If for example a solid plate separates two different liquids (a lower liquid and an upper liquid), then the angle of incidence, for stimulating the same Lamb mode, is unchanged, whereas the inhomogeneity differs compared with the situation where both liquids are equal to each other. If for a given upper liquid and for a given Lamb mode this inhomogeneity difference is plotted as a function of the impedance of the lower liquid, then a linear dependence is noticed. It is noticed that all physical liquids follow this linear tendency. Furthermore, this linearity depends on the generated mode and also on the properties of the solid plate. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

9:15

2aUW6. Born approximation for scattering by evanescent waves: Comparison with exact scattering by an infinite fluid cylinder. Philip L. Marston (Phys. Dept., Washington State Univ., Pullman, WA 99164-2814)

In some situations, evanescent waves can be an important component of the acoustic field within the sea bottom. For this reason (as well as to advance the understanding of scattering processes) it can be helpful to examine the modifications to scattering theory resulting from evanescence. Modifications to ray theory were examined in a prior approximation [P. L. Marston, J. Acoust. Soc. Am. **113**, 2320 (2003)]. The new research concerns the modifications to the low-frequency Born approximation and confirmation by comparison with the exact two-dimensional scattering by a fluid cylinder. In the case of a circular cylinder having the same density as the surroundings but having a compressibility contrast with the surroundings, the Born approximation with a nonevanescent incident wave gives only monopole scattering. When the cylinder has a density contrast and the same compressibility as the surroundings the regular Born approximation gives only dipole scattering (with the dipole oriented along to the incident wavevector). In both cases when the Born approximation is modified to include the evanescence of the incident wave, an additional dipole scattering term is evident. In each case the new dipole is oriented along to the decay axis of the evanescent wave. [Research supported by ONR.]

9:30

2aUW7. Fluid-filled spheres: Theoretical and measured scattering response. David M. Deveau (Naval Undersea Warfare Ctr., 801 Clematis St., West Palm Beach, FL 33401, deveau@wpb.nuwc.navy.mil)

Investigation into the scattering nature of surfaces or other physical objects often requires the use of measurement systems which cannot always be well controlled. This lack of control can be compensated for by calibrating the resulting measurements against a known target. While these targets can be any object, the goal is to use a target that has a stable scattering response and is independent of angle. The ideal shape is that of a sphere, but even this can be improved with the addition of internal fluids that focus and temperature stabilize the scattering response. A scatter response model of a sphere has been developed and used to design four thin-walled spheres, each with a different diameter and filled with a focusing fill fluid (Fluorolube). One pair of spheres was measured in an ocean environment, while the second pair was tested in a controlled test tank from 5–50 kHz but using shorter continuous wave pulses. While the ocean-measured spheres closely matched the model, the test-tank measurements showed a marked difference from the model. Changes to the model will be explored to determine if theoretical minimums for pulse length are insufficient for targets of this density or focusing capability.

9:45

2aUW8. Scattering from bubbles rising in a vertical line—Comparison between theory and experiments. Aneta Nikolovska, Andrew Ooi (Dept. of Mech. and Manufacturing Eng., The Univ. of Melbourne, Parkville, VIC 3010, Melbourne, Australia), and Richard Manasseh (CSIRO Thermal and Fluids Eng., Highett, VIC 3190, Melbourne, Australia)

In this paper, the experimental data from the sound field around similar sized air bubbles rising in a vertical chain are analyzed. The data reveal a strong anisotropy in the acoustic field. The transition from individual to continuum behavior in a bubble chain is not correctly described with classical theory especially when the bubbles are uniformly sized, discretely populated, and the frequency of interest is close to the natural frequency of the individual bubbles. Single compact scatterers initiated at frequencies near their resonant frequency in isolation act preliminarily as monopole sources, amplifying the local pressure field by a factor of order $1/(k^*a)$, a being the scattering radius and k the wave number at the resonant frequency in the surrounding material [Tolstoy, 1986]. A laboratory investigation used air bubbles in fresh water and varied the bubble sizes and separation by carefully controlling bubble production rates. A linear coupled equation method was developed to explain the result. The model reproduced the acoustic pressure anisotropy along the chain and the change in pulse waveform along the chain. The results suggest that the enhancement of sound intensity along the chain can be explained by bubbles acting as resonant scatterers retransmitting the acoustic energy.

10:00–10:15 Break

10:15

2aUW9. Development of a submersible, baffled impedance tube for characterization of bubbly liquids. Jed C. Wilbur, Jason Holmes, Michael Canney, William M. Carrey, and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215)

The ability to measure the acoustic properties of bubbly liquids *in situ* has long been a goal of the underwater acoustics community. An apparatus for characterizing these properties in the 1–16-kHz frequency range capable of submerged deployment in a controlled environment (large tank or test pond) has been developed. A water-filled, 2.5-cm-thick walled, baffled, stainless steel impedance tube was used to measure the complex reflection coefficient at the opening. The classical result for the radiation impedance of a baffled, open tube is verified, as is that for the radiated sound field along the tube axis. The test tank is then filled with air bubbles (order 500 micron mean radius and 0.1% vf) and the measurement repeated at the water/bubbly liquid interface. The acoustic properties of the bubbly liquid were obtained as a function of frequency by an inversion process using classical theory. Because of the high attenuation in bubbly liquids, the finite-sized baffle (61-cm outer diameter) is treated as being infinite in extent. The measured sound speeds and attenuations are compared with established theory. [Work supported by U.S. Navy ONR.]

10:30

2aUW10. Optimal passive source localization in a fluctuating ocean waveguide based on an analytic model for the mean field and covariance. Purnima Ratilal, Ioannis Bertsatos, Tianrun Chen, Michele Zanolin, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, purnima@mit.edu)

The ocean acts as an enormously complex channel for signal transmission. It is characterized by temporal fluctuations and spatial variabilities that often lead to significant randomization of the measured field. Signal processing techniques must be derived from fundamental physical models to attain statistical optimality in such an environment. Here we present such an analysis with a new analytic model [Ratilal and Makris, J. Acoust. Soc. Am. **114**, 2428 (2003)] for the mean field and its covariance after propagation through a fluctuating ocean waveguide based on the first principles of waveguide scattering theory. The model is advantageous because it includes the primary physical effects of attenuation, dispersion, and coupling of modal energy due to multiple forward scattering in a convenient and intuitive form that is well suited to analytic manipulations. An example will be presented for passive source localization in range and depth in a fluctuating ocean waveguide by nonlinear matched field inversion. The necessary conditions on sample size and signal-to-noise ratio (SNR) for unbiased and minimum variance (optimal) estimates of source postion are then derived using asymptotic statistical inference theory [Naftali and Makris, J. Acoust. Soc. Am. **110**, 1917–1930 (2001)]. Degradation in performance by environmental uncertainties are quantified within this statistical framework.

10:45

2aUW11. Some interesting features in pulse propagation for environments with structured velocity profiles. R. Field and M. Werby (NRL Code 7181, Stennis Space Ctr., MS 39529)

On a diurnal time scale the velocity profiles in some shallow water environments can change enough to modify signal detection features. In particular, if a vertical array is used to collect data, then contour plots of

the detected signals relative to time and vertical distance can reflect features of the environment that are readily interpreted and therefore reveals some details of interest. Two features are investigated here. One pertains to surface ducting, which may be present over some limited time scale on a daily basis. The other feature is a rather pronounced variation in bottom ducting diurnally. At suitable frequencies (1 KH–5 KH) contour plots indicate discrete time arrivals due to sound waves trapped in the surface duct that reflect fast early (precursor) group arrivals. This effect has been studied earlier by H. DeFerrari and can be useful in extracting environmental features. Broad bottom ducts manifest themselves as broad group fronts characterized by almost vertical initial arrivals. This pattern varies on a daily basis. Thus these contours offer a simple method for diurnal duct variations and ought to be of interest for passive detection.

11:00

2aUW12. Recovery and digitization of archival analog ocean acoustic data. Jack Shooter (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029) and Roy D. Gaul (BlueSea Corp.)

In October 2002 the Office of Naval Research (ONR) sponsored a Convocation that reviewed the ocean acoustic studies done from 1967 to 1992 under the U.S. Navy's Long Range Acoustic Propagation Project (LRAPP). Beginning in 1972 LRAPP fielded self-contained assemblies of vertically distributed hydrophones as part of environmental acoustic exercises in a variety of oceanic regions. Analog signals were recorded in a submerged buoy on multi-channel magnetic tape. Presentation of results from a 1975 measurements exercise in the Northeast Pacific stimulated interest in recovering and digitizing the 10-day dataset from 13 hydrophones. ONR sponsored a pilot project to demonstrate the feasibility of recovering the data and setting up digital files while preserving absolute accuracy. The original magnetic tape was sticky. It had to be heat treated and cleaned before playback. A spectrum from the original analysis for near-field passage of a ship was used to validate data recovery in the range of 10–500 Hz. A final objective is to produce calibrated time series so that the digital dataset can be made available for general use. [Work supported by ONR.]

11:15

2aUW13. Passive acoustic localization for Sciaenid habitat in coastal water of Taiwan. Chang Tu, Ruey-Chang Wei (Inst. of Undersea Technol., Natl. Sun Yat-sen Univ., 70, Lien-hai Rd., Kaohsiung, Taiwan, rcwei@mail.nsysu.edu.tw), and Hsiang-Chih Chan (Natl. Taiwan Univ., Taipei 804, Taiwan)

There are many Sciaenid species found in coastal water of Taiwan, and most of them can generate sound in the spawning season. However, due to overfishing, the populations of these high economic value fishes have been greatly decreased. To study and protect Sciaenid, whose habitat should be identified, seasonal protection zones should be set up. In this study, easy and low-cost measure of using passive sonar is proposed to map the habitat in the field. Measurement and analysis of Sciaenid sound were performed to investigate its acoustic characteristic, and frequency range was found to be 350 to 1500 Hz. The statistical distribution of sound levels can be built up by enough field measurement; therefore source levels can be defined in this method. In addition, the transmission loss of in-site shallow water was studied to compute the real decaying factor ($TL = n * \log R$) of the survey site. By using three sonobuoys, with known source level distribution and transmission loss, the distance of possible habitat can then be inverted. Through enough measurements, the map of coastal habitat of Sciaenid can be generated with efficiency and accuracy.

11:30

2aUW14. Development of net cage acoustic alarm system. Shih-Wei Hong, and Ruey-Chang Wei (Inst. of Undersea Technol., Natl. Sun Yat-sen Univ., 70, Lien-hai Rd., Kaohsiung, Taiwan, rcwei@mail.nsysu.edu.tw)

In recent years, the fishery production has been drastically decreased in Taiwan, mainly due to overfishing and coast pollution; therefore, fishermen and corporations are encouraged by government to invest in ocean net cage aquaculture. However, the high-price fishes in the net cage are often coveted, so incidences of fish stealing and net cage breaking were found occasionally, which cause great economical loss. Security guards or a visual monitoring system has limited effect, especially in the night when these intrusions occur. This study is based on acoustic measure to build a net cage alarm system, which includes the sonobuoy and monitor station on land. The sonobuoy is a passive sonar that collects the sounds near the net cage and transmits the suspected signal to the monitor station. The signals are analyzed by the control program on the personal computer in the monitor station, and the alarms at different stages could be activated by the sound levels and durations of the analyzed data. To insure long hours of surveillance, a solar panel is applied to charge the battery, and a photodetector is used to activate the system.

TUESDAY MORNING, 25 MAY 2004 CONFERENCE ROOM L, 8:00 TO 8:50 A.M.

Meeting of the Standards Committee Plenary Group

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meetings for: ISO/TC 43 Acoustics ISO/TC 43/SC 1 Noise, and IEC/TC 29 Electroacoustics

The meeting of the Standards Committee Plenary Group will precede the meetings of the Accredited Standards Committees S1, S2, S3, and S12, which are scheduled to take place later on the same day. Discussion at the Standards Committee Plenary Group meeting will consist of national items relevant to all four S Committees.

The ANSI-Accredited U.S. Technical Advisory Group (TAGs) for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics, whose membership consists of members of S1 and S3, and other persons not necessarily members of these Committees, will meet during the Standards Plenary meeting. The ANSI-Accredited US Technical Advisory Group (TAG) for ISO/TC 43/SC 1 Noise, whose membership consists of the members of S12 and other persons not necessarily members of S12, will meet as well. *The reports of the Chairs*

of these TAGs will not be presented at any other S Committee meeting. There will be a report on the interface of S1 and S3 activities with those of ISO/TC 43 and IEC/TC 29 including plans for future meetings of ISO/TC 43 and IEC/TC 29.

Members of S2 Mechanical Vibration and Shock (and U.S. TAG for ISO/TC 108 and four of its Subcommittees, (SC2, SC3, SC5, and SC6) are also encouraged to attend the Standards Committee Plenary Group meeting even though the S2 meeting will take place later in the day.

TUESDAY MORNING, 25 MAY 2004 CONFERENCE ROOM L, 9:00 TO 10:00 A.M.

Meeting of Accredited Standards Committee (ASC) S1 Acoustics

Accredited Standards Committee S1 on Acoustics. Working group chairs will report on the status of standards currently under development in the areas of physical acoustics, electroacoustics, sonics, ultrasonics, and underwater sound, etc. Consideration will be given to new standards that might be needed over the next few years. Open discussion of committee reports is encouraged.

People interested in attending the meeting of the TAGs for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics, take note—those meetings will be held in conjunction with the Standards Plenary meeting at 8:00 a.m. on Tuesday, 25 May 2004.

Scope of S1: Standards, specifications, methods of measurement and test, and terminology in the field of physical acoustics, including architectural acoustics, electroacoustics, sonis and ultrasonics, and underwater sound, but excluding those aspects which pertain to biological safety, tolerance and comfort.

TUESDAY MORNING, 25 MAY 2004 CONFERENCE ROOM L, 10:30 TO 12:00 NOON

Meeting of Accredited Standards Committee (ASC) S12 Noise

Accredited Standards Committee S12 on Noise. Working group chairs will report on the status of noise standards currently under development. Consideration will be given to new standards that might be needed over the next few years. Open discussion of committee reports is encouraged.

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.

TUESDAY AFTERNOON, 25 MAY 2004 ROYAL BALLROOM A, 2:00 TO 3:25 P.M.

Session 2pAAa

Architectural Acoustics: Methods to Quantify Opera House Acoustics III: Simulations and Auralizations

Jin Yong Jeon, Chair
School of Architectural Engineering, Hangyang University, 12 Haengding-dong, Seongdong-gu, Seoul 133-79, Korea

Chair's Introduction—2:00

Invited Papers

2:05

2pAAa1. An experimental comparative study of 20 Italian opera houses: Measurement techniques. Angelo Farina, Enrico Armelloni, and Paolo Martignon (Indust. Eng. Dept., Univ. of Parma, Via delle Scienze 181/A, 43100 Parma, Italy)

For "acoustical photography" we mean a set of measured impulse responses, which enable us to "listen" at the measured room by means of advanced auralization methods. Once these data sets have been measured, they can be employed in two different ways: objective analysis and listening test. In fact, it is possible to compute dozens of acoustical objective parameters, describing the temporal texture, the spatial effect and the frequency-domain coloring of each opera house. On the other hand, by means of the

auralization technique, it becomes easy to conduct listening experiments with human subjects. This paper focuses principally on the development and specification of the measurement technique, which is the topic assigned to the research unit of Parma, to which the authors belong. It describes the hardware equipment, the software, the electro-acoustic transducers (microphones and loudspeakers), the measurement positions, the system for automatic displacement of the microphones and the conditions of the room during the measurements. Experimental results are reported about a couple of opera houses which were employed for testing the measurement procedure and showing the benefits of the new method against the previously employed ones.

2:25

2pAAa2. Auditorium acoustics evaluation based on simulated impulse response. Shuoxian Wu, Hongwei Wang, and Yuezhe Zhao (Dept. of Architecture, South China Univ. of Technol., Guangzho 510640, China)

The impulse responses and other acoustical parameters of Huangpu Teenager Palace in Guangzhou were measured. Meanwhile, the acoustical simulation and auralization based on software ODEON were also made. The comparison between the parameters based on computer simulation and measuring is given. This case study shows that auralization technique based on computer simulation can be used for predicting the acoustical quality of a hall at its design stage.

2:45

2pAAa3. Calculating auditorium acoustic parameters from measured binaural impulse response. Yuezhe Zhao and Shuoxian Wu (Dept. of Architecture, South China Univ. of Technol., Guangzho 510640, China)

The differences between Chinese traditional music and western music will cause some specific acoustics requirements for the auditoriums where national music programs are played. There are varieties of coupled spaces in buildings. These coupled spaces have some special acoustical characteristics. As a starting point of these researches, the binaural impulse responses and other acoustical parameters were measured and calculated for Guangzhou Southern Theatre where Chinese traditional music and Guangdong folk music are usually performed. The measurement results were analyzed in detail herein.

3:05

2pAAa4. Acoustical measurements in a 1:10 scale model of a multi-purpose hall. Jin Yong Jeon and Jong Kwan Ryu (School of Architectural Eng., Hanyang Univ., Seoul 133-791, Korea)

Measurements were made in a 1:10 scale model of a 1500-seat multi-purpose hall to investigate the acoustic properties of its opera mode. The length from the stage front to the rear wall is 33 m and its maximum width is 28 m with a volume of 14300 m^3. Reflecting surfaces were included to scatter early energy from the stage area onto the audience areas. Several settings of the sources were used on the platform and in the pit in order to determine how a lack of brilliance caused by diffraction occurred in the stalls audience. The advantages of the curved ceiling surface in front of the proscenium above the orchestra pit were exploited. In addition, the effect of some diffusing treatments on surfaces adjacent to the pit and on the side walls was also investigated.

TUESDAY AFTERNOON, 25 MAY 2004 ROYAL BALLROOM A, 3:30 TO 5:15 P.M.

Session 2pAAb

Architectural Acoustics: Methods to Quantify Opera House Acoustics IV: Measurements

Richard H. Campbell, Chair

Bang Campbell Associates, 3 Water Street, Woods Hole, Massachusetts 02543

Chair's Introduction—3:30

Invited Papers

3:35

2pAAb1. Methods to quantify opera houses acoustics—Changes to why, what, and how? Robert W Harris (Arup Acoust., Parkin House, 8 St. Thomas St., Winchester SO23 9HE, UK, rob.harris@arup.com)

The historical and current reasons for quantifying opera house acoustics (the why?) are reviewed. The emerging shift from quantitiative assessment (by measurement) to qualitative assessment (by auralization) is discussed. The key characteristics of the acoustics of opera houses (the what?) are then considered. Assessment of a theatre with an audience present is suggested to be more important than unoccupied measurements. This has implications for the development of assessment methods. The suitability of classical room acoustics parameters—mostly derived from concert hall studies—is debated (the how?). Possible new assessment techniques are emerging. Characteristics not addressed by the classical parameters include the early (ensemble) and late (room) support provided to the singers and the onset threshold of image shifting of the sound of instruments within the orchestra pit. The assessment of loudness balance between the stage and the pit requires refinement and development. A difficulty with the assessment and comparison of opera houses is the effect of the stage settings on the acoustics.

3:55

2pAAb2. Design of a new stage shell for the Stadthaus in Winterthur, Switzerland. Eckhard Kahle (Kahle Acoust., 156 Ave. du Parc, B-1190 Brussels, Belgium, kahle@skynet.be) and Brian F. G. Katz (LIMSI-CNRS, F-91403 Orsay, France)

A new stage shell was installed in the Stadthaus, a 700-seat concert hall in Winterthur which was originally designed by architect Gottfried Semper. Musicians of the local symphony orchestra had been complaining about poor on-stage listening conditions in the stage house that was later added to the room. The small stage house (width 12 m, depth 7 m, height 7 m) was overly loud and reverberant. The design of the new stage shell will be presented and the acoustic results, both objective and subjective, will be presented. Comments by the musicians and audience members show that a stage shell must provide more than just efficient projection into the audience area. Absorption placed at the rear of the stage is paramount for obtaining proper orchestral balance, good on-stage hearing condition and clarity of sound.

Contributed Papers

4:15

2pAAb3. Measurements of the self-to-other ratio in an opera chorus in performance. Sten Ternstrom (Dept. of Speech, Music & Hearing, KTH, Stockholm, sten@speech.kth.se) and Pamela Davis (Univ. of Sydney, Sydney, Australia)

The objective of this study was to obtain measurements of the self-to-other ratio (SOR) in the opera chorus on stage. Four members of the opera chorus of Opera Australia volunteered to carry wireless binaural microphones during a dress rehearsal of Verdi's *The Masked Ball*, in the Sydney Opera House. Conditions were those of actual performance in costume, with soloists and orchestra. Hence, the recordings are very realistic, if acoustically not so stringently controlled. The SOR was estimated from the recordings using M/S matrixing, complemented with L–R correlation by adaptive filtering. The results indicate how much louder the singers heard themselves relative to the rest of the ensemble, including the orchestra. The SOR varied with score and on-stage formation, but was 4–8 dB higher than in ordinary choirs (not opera), as might be expected from the podium acoustics. While each opera chorus artist can hear his or her own voice very well, they often hear less of the rest of the chorus and very little of the orchestra. This was borne out in informal listening; the orchestra often became inaudible once the choir entered its louder passages. [Work supported by the Australian Research Council under the SPIRT program.]

4:30

2pAAb4. The measurements technique of balance in opera houses: A study of the sources involved. Linda Parati, Roberto Pompoli, and Nicola Prodi (Dept. of Eng., Univ. of Ferrara, 44100 Ferrara, Italy, lparati@ing.unife.it)

The study of the acoustical balance between the singer and the orchestra by means of room acoustical measurements has shown that the directional characteristics of the source on the stage are important. This investigation compares the performance of two directional and one omnidirectional loudspeakers in emulating a soprano voice. Directivity measurements of a soprano singer were carried out in an anechoic chamber and used as a basis for comparison in room acoustic simulations of the Royal Theatre of Copenhagen, the Ankara Congress and Cultural Center, and the Alberta Jubilee Auditorium. In particular, the balance measurement was simulated and the performance of the different sources with respect to the soprano was assessed.

4:45

2pAAb5. Acoustical measurements in ancient Roman theatres. Andrea Farnetani, Patrizio Fausti, Roberto Pompoli, and Nicola Prodi (Eng. Dept., Univ. of Ferrara, Via Saragat 1, 44100 Ferrara, Italy, afarnetani@ing.unife.it)

The Greek and Roman theatres are among the most precious and spectacular items of cultural heritage in the Mediterranean countries. The theatres are famous not only for their impressive architecture, but also for the acoustic qualities. For this reason it is important to consider these theatres as an acoustical heritage and to study their sound field. Within the activities of the ERATO (identification Evaluation and Revival of the Acoustical heritage of ancient Theatres and Odea) project, acoustical measurements were taken in well-preserved ancient Roman theatres at Aspendos (Turkey) and Jerash (Jordan). Roman theatres have an impressive stage building that forms a back wall in the orchestra area, and it was found that, from the analysis of the acoustical parameters, the reverberation time (e.g., 1.7 s at middle frequencies in the theatre of Aspendos) is quite long compared not only with other open-space theatres but also with closed spaces. Contrary to modern halls the clarity is high and this fact, together with a low sound level in most of the seats, gives the sound field a unique character.

5:00

2pAAb6. Acoustical measurement of the Salt Lake Mormon Tabernacle. Sarah Rollins and Timothy W. Leishman (N283 Eyring Sci. Ctr., Brigham Young Univ., Provo, UT 84602)

An acoustical survey of the Salt Lake Mormon Tabernacle has been performed to assess the behavior of the hall in its current state. The tabernacle is a well-known historical building with a large elongated dome ceiling. This paper discusses the measurements used to characterize the hall. Several parameters derived from omnidirectional, directional, and binaural impulse response measurements are presented. Color maps of the parameters over audience seating areas reveal their spatial variations. These maps and the statistical properties of the parameters aid in clarifying the acoustical characteristics and anomalies of the hall.

Session 2pAAc

Architectural Acoustics: Theaters for Drama Performance: Another Two Decades (1984–2004) (Poster Session)

Gregory Miller, Chair
The Talaske Group, 105 N. Oak Park Drive, Oak Park, Illinois 60301

Contributed Papers

All posters will be on display from 1:30 p.m. to 5:30 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:30 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 5:30 p.m.

2pAAc1. Denver Center for the Performing Arts, Bonfils Theatre Complex. R. Lawrence Kirkegaard (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc2. Portland Performing Arts Center. R. Lawrence Kirkegaard (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc3. Broward Center for the Performing Arts. R. Lawrence Kirkegaard (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc4. Guthrie Theatre. Joseph Myers (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc5. Oregon Shakespeare Festival. Joseph Myers (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc6. Blumenthal Center for the Performing Arts. R. Lawrence Kirkegaard (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc7. University of Cincinnati, College Conservatory of Music. Joseph Kirkegaard (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc8. Berkeley Repertory RODA Theater. Jason R. Duty and David R. Schwind (Charles M. Salter Assoc., Inc., 130 Sutter St., Ste. 500, San Francisco, CA 94104, acoustics@cmsalter.com)

2pAAc9. Aronoff Center. Dawn Schuette (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc10. University of Alabama at Birmingham, Alys Robinson Stephens Performing Arts Center. Joseph Myers (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc11. North Shore Center for the Performing Arts. Dawn Schuette (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc12. Milton Keynes Theatre, Milton Keynes, UK. Rafal J. Orlowski and Rob Harris (Arup Acoust., St. Giles Hall, Pound Hill, Cambridge CB3 0AE, UK)

2pAAc13. Babson College Campus Center. Joseph Myers (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc14. Barbican Centre, Royal Shakespeare Company Theatre. Carl Giegold (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc15. Overture Center. Joseph Myers (Kirkegaard Assoc., 801 W. Adams St., 8th Fl., Chicago, IL 60607, acoustics@kirkegaard.com)

2pAAc16. Hampstead Theatre, London. Ian Thompson (Arup Acoust., St. Giles Hall, Pound Hill, Cambridge CB3 0AE, UK)

2pAAc17. Founders Theater. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc18. South Dakota State University Theater. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc19. Courtyard Theatre. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc20. Lucille Little Theater Transylvania University. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc21. Ohio Theater. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc22. Williams College Theatre. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc23. Regent Theatre. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc24. Southern Theatre. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc25. Stanley Theatre. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc26. University School Theater. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc27. Yamaguchi Center for Arts and Media, Studio A. Makoto Ino and Keiji Oguchi (Nagata Acoustics, Inc., Hongo-Segawa Bldg., 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan)

2pAAc28. Playwright's Horizon. L. Gerald Marshall (Marshall KMK, 59 S. Greeley Ave., Chappaqua, NY 10514) and K. Anthony Hoover (Cavanaugh Tocci Assoc., Inc., Sudbury, MA 01776)

2pAAc29. Four Rivers Center for the Performing Arts. Katharine Sawicki (Artec Consultants, Inc., 114 W. 26th St., New York, NY 10001, ks@artec-usa.com)

2pAAc30. Toyota City Noh Theater. Keiji Oguchi and Katsuji Naniwa (Nagata Acoustics, Inc., Hongo-Segawa Bldg., 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan)

2pAAc31. Curtis M. Phillips Center for the Performing Arts, University of Florida, Gainesville. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2pAAc32. Kitakyusyu Performing Arts Center, Theatre. Keiji Oguchi and Katsuji Naniwa (Nagata Acoustics, Inc., Hongo-Segawa Bldg., 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan)

2pAAc33. Bromsgrove Arts Center. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc34. Ruth Shapiro Theater, Carl and Ruth Shapiro Student Center, Brandeis University. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2pAAc35. Theater 315—Salvation Army Times Square. Ted Pyper (Artec Consultants, Inc., 114 W. 26th St., New York, NY 10001, tap@artec-usa.com)

2pAAc36. Skirball Center for the Performing Arts. Ted Pyper (Artec Consultants, Inc., 114 W. 26th St., New York, NY 10001, tap@artec-usa.com)

2pAAc37. The Chan Centre for the Performing Arts—Telus Studio Theatre. Katharine Sawicki (Artec Consultants, Inc., 114 W. 26th St., New York, NY 10001, ks@artec-usa.com)

2pAAc38. Roundhouse—Royal Shakespeare Company. Alicia Koledin (Acoustic Dimensions, 145 Huguenot St., Ste. 406, New Rochelle, NY 10801, info@acousticdimensions.com)

2pAAc39. Art Sphere Theater. Yasuhisa Toyota and Katsuji Naniwa (Nagata Acoustics, Inc., U.S. Office, 1610 Broadway, Ste. A, Santa Monica, CA 90404)

2pAAc40. Interlochen Center for the Arts—Harvey Theatre. Todd L. Brooks (Artec Consultants, Inc., 114 W. 26th St., New York, NY 10001, tlb@artec-usa.com)

2pAAc41. Fox Cities Performing Arts Center. Todd L. Brooks (Artec Consultants, Inc., 114 W. 26th St., New York, NY 10001, tlb@artec-usa.com)

2pAAc42. Performing Arts Center, The Groton School. Benjamin Markham (Acentech, Inc., 33 Moulton St., Cambridge, MA 02138)

2pAAc43. Tokyo Metropolitan Art Space's Theater. Satoru Ikeda, Keiji Oguchi, Makoto Ino, and Minoru Nagata (Nagata Acoustics, Inc., Hongo-Segawa Bldg., 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan)

2pAAc44. Shizuoka Arts Theatre. Satoru Ikeda, Chiaki Ishiwata, and Motoo Komoda (Nagata Acoustics, Inc., Hongo-Segawa Bldg., 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan)

2pAAc45. Ishikawa Ongakudo's Hougaku Hall. Satoru Ikeda, Keiji Oguchi, and Masaya Uchida (Nagata Acoustics, Inc., Hongo-Segawa Bldg., 3F, 2-35-10 Hongo Bunkyo-ku, Tokyo 113-0033, Japan)

2pAAc46. Princess of Wales Theatre. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada, jokeefe@aercoustics.com)

2pAAc47. Evergreen Cultural Centre. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc48. Ruth Seaton James Centre for the Performing Arts. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc49. A. E. Rawlinson Centre for Arts. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc50. Hummingbird Centre for the Performing Arts. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc51. Theatre Aquarius. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc52. Victoria Hall. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc53. Teatro Sesc Santo Amaro, Sao Paulo, Brazil. Jose Augusto Nepomuceno (Acustica & Sonica, Rua Fradique Coutinho, 955 sala 01 05416-011, Sao Paulo SP, Brazil, janepomuceno@yahoo.com)

2pAAc54. Teatro de Araras, Araras, Brazil. Jose Augusto Nepomuceno (Acustica & Sonica, Rua Fradique Coutinho, 955 sala 01 05416-011 Sao Paulo SP, Brazil, janepomuceno@yahoo.com)

2pAAc55. Teatro Sesc Santana, Sao Paulo, Brazil. Jose Augusto Nepomuceno (Acustica & Sonica, Rua Fradique Coutinho, 955 sala 01 05416-011, Sao Paulo SP, Brazil, janepomuceno@yahoo.com)

2pAAc56. Theatro Sao Pedro, Sao Paulo, Brazil. Jose Augusto Nepomuceno (Acustica & Sonica, Rua Fradique Coutinho, 955 sala 01 05416-011, Sao Paulo SP, Brazil, janepomuceno@yahoo.com)

2pAAc57. Showplace Performance Centre. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc58. The Esplanade. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc59. Teatros dos CEUs, Sao Paulo, Brazil. Jose Augusto Nepomuceno (Acustica & Sonica, Rua Fradique Coutinho, 955 sala 01 05416-011, Sao Paulo SP, Brazil, janepomuceno@yahoo.com)

2pAAc60. Young Centre for the Performing Arts. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc61. Teatro Alfa, Sao Paulo, Brazil. Jose Augusto Nepomuceno (Acustica & Sonica, Rua Fradique Coutinho, 955 sala 01 05416-011, Sao Paulo SP, Brazil, janepomuceno@yahoo.com)

2pAAc62. Chinese Cultural Centre in Toronto. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc63. Queen Elizabeth Theatre. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

2pAAc64. New Amsterdam Theater. Paul Scarbrough (Akustiks, 11 N. Main St., Norwalk, CT 06854, pscarbrough@akustiks.net) and David Greenberg (Creative Acoustics, Westport, CT 06880-6134[a])

[a]Paul Scarbrough was principal-in-charge and David Greenberg was the project manager for Jaffe Holden Scarbrough Acoustics.

2pAAc65. New Victory Theater. Paul Scarbrough and Anthony Nittoli (Akustiks, 11 N. Main St., Norwalk, CT 06854, pscarbrough@akustiks.net)[a]

[a]Paul Scarbrough was the principal-in-charge for Jaffe Holden Scarbrough Acoustics. Anthony Nittoli was senior project manager for Pro-Mix.

2pAAc66. Isabel Bader Theatre. John O'Keefe (50 Ronson Dr., Ste. 165, Toronto, ON, Canada)

Session 2pAAd

Architectural Acoustics and Engineering Acoustics: Bell Labs and Acoustics

Gary W. Elko, Cochair
Avaya Labs, Basking Ridge, New Jersey 07920

Neil A. Shaw, Cochair
Menlo Scientific Acoustics, Inc., P.O. Box 1610, Topanga, California 90290-1610

James E. West, Cochair
Department of Electrical and Computer Engineering, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218-2686

Chair's Introduction—1:00

Invited Papers

1:05

2pAAd1. Communication acoustics in Bell Labs. J. L. Flanagan (Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854-8088)

Communication aoustics has been a central theme in Bell Labs research since its inception. Telecommunication serves human information exchange. And, humans favor spoken language as a principal mode. The atmospheric medium typically provides the link between articulation and hearing. Creation, control and detection of sound, and the human's facility for generation and perception are basic ingredients of telecommunication. Electronics technology of the 1920s ushered in great advances in communication at a distance, a strong economical impetus being to overcome bandwidth limitations of wireline and cable. Early research established criteria for speech transmission with high quality and intelligibility. These insights supported exploration of means for efficient transmission—obtaining the greatest amount of speech information over a given bandwidth. Transoceanic communication was initiated by undersea cables for telegraphy. But these long cables exhibited very limited bandwidth (order of few hundred Hz). The challenge of sending voice across the oceans spawned perhaps the best known speech compression technique of history—the Vocoder, which parametrized the signal for transmission in about 300 Hz bandwidth, one-tenth that required for the typical waveform channel. Quality and intelligibility were grave issues (and they still are). At the same time parametric representation offered possibilities for encryption and privacy inside a traditional voice bandwidth. Confidential conversations between Roosevelt and Churchill during World War II were carried over high-frequency radio by an encrypted vocoder system known as Sigsaly. Major engineering advances in the late 1940s and early 1950s moved telecommunications into a new regime—digital technology. These key advances were at least three: *(i)* new understanding of time-discrete (sampled) representation of signals, *(ii)* digital computation (especially binary based), and *(iii)* evolving capabilities in microelectronics that ultimately provided circuits of enormous complexity with low cost and power. Digital transmission (as exemplified in pulse code modulation—PCM, and its many derivatives) became a telecommunication mainstay, along with switches to control and route information in digital form. Concomitantly, storage means for digital information advanced, providing another impetus for speech compression. More and more, humans saw the need to exchange speech information with machines, as well as with other humans. Human-machine speech communication came to full stride in the early 1990s, and now has expanded to multimodal domains that begin to support enhanced naturalness, using contemporaneous sight, sound and touch signaling. Packet transmission is supplanting circuit switching, and voice and video are commonly being carried by Internet protocol.

1:25

2pAAd2. Digital signal processing at Bell Labs—Foundations for speech and acoustics research. Lawrence R. Rabiner (CAIP Ctr., Rutgers Univ., 96 Frelinghuysen Rd., Piscataway, NJ 08854-8088, lrr@caip.rutgers.edu)

Digital signal processing (DSP) is a fundamental tool for much of the research that has been carried out of Bell Labs in the areas of speech and acoustics research. The fundamental bases for DSP include the sampling theorem of Nyquist, the method for digitization of analog signals by Shannon *et al.*, methods of spectral analysis by Tukey, the cepstrum by Bogert *et al.*, and the FFT by Tukey (and Cooley of IBM). Essentially all of these early foundations of DSP came out of the Bell Labs Research Lab in the 1930s, 1940s, 1950s, and 1960s. This fundamental research was motivated by fundamental applications (mainly in the areas of speech, sonar, and acoustics) that led to novel design methods for digital filters (Kaiser, Golden, Rabiner, Schafer), spectrum analysis methods (Rabiner, Schafer, Allen, Crochiere), fast convolution methods based on the FFT (Helms, Bergland), and advanced digital systems used to implement telephony channel banks (Jackson, McDonald, Freeny, Tewksbury). This talk summarizes the key contributions to DSP made at Bell Labs, and illustrates how DSP was utilized in the areas of speech and acoustics research. It also shows the vast, worldwide impact of this DSP research on modern consumer electronics.

1:45

2pAAd3. Speech coding research at Bell Laboratories. Bishnu S. Atal (Dept. of Elec. Eng., Univ. of Washington, Seattle, WA 98195 and Atal-Res., Mukilteo, WA 98275)

The field of speech coding is now over 70 years old. It started from the desire to transmit voice signals over telegraph cables. The availability of digital computers in the mid 1960s made it possible to test complex speech coding algorithms rapidly. The introduction of linear predictive coding (LPC) started a new era in speech coding. The fundamental philosophy of speech coding went through a major shift, resulting in a new generation of low bit rate speech coders, such as multi-pulse and code-excited LPC. The semiconductor revolution produced faster and faster DSP chips and made linear predictive coding practical. Code-excited LPC has become the method of choice for low bit rate speech coding applications and is used in most voice transmission standards for cell phones. Digital speech communication is rapidly evolving from circuit-switched to packet-switched networks to provide integrated transmission of voice, data, and video signals. The new communication environment is also moving the focus of speech coding research from compression to low cost, reliable, and secure transmission of voice signals on digital networks, and provides the motivation for creating a new class of speech coders suitable for future applications.

2:05

2pAAd4. Hands-free human-machine interaction with voice. B. H. Juang (Georgia Inst. of Technol., Atlanta, GA 30332)

Voice is natural communication interface between a human and a machine. The machine, when placed in today's communication networks, may be configured to provide automation to save substantial operating cost, as demonstrated in AT&T's VRCP (Voice Recognition Call Processing), or to facilitate intelligent services, such as virtual personal assistants, to enhance individual productivity. These intelligent services often need to be accessible anytime, anywhere (e.g., in cars when the user is in a hands-busy-eyes-busy situation or during meetings where constantly talking to a microphone is either undersirable or impossible), and thus call for advanced signal processing and automatic speech recognition techniques which support what we call ''hands-free'' human-machine communication. These techniques entail a broad spectrum of technical ideas, ranging from use of directional microphones and acoustic echo cancellatiion to robust speech recognition. In this talk, we highlight a number of key techniques that were developed for hands-free human-machine communication in the mid-1990s after Bell Labs became a unit of Lucent Technologies. A video clip will be played to demonstrate the accomplishement.

2:25

2pAAd5. On my association with Bell Labs. M. Mohan Sondhi (Avaya Labs, 233 Mount Airy Rd., Basking Ridge, NJ 07920)

I joined the Acoustics Research department at Bell Labs in 1962, just eight days before AT&T launched the first communications satellite, Telstar. During the 39 years between 1962 and my retirement in 2001, I worked on several problems related in one way or another to the processing of speech signals. Schroeder and Flanagan are presenting talks from a broad perspective in this session, so I will confine this talk to just my own contributions and collaborations for some of the topics on which I worked, e.g., echo cancellation, inverse problems in acoustics, speech analysis, synthesis, and recognition. I will tell you about one of these contributions that fortunately turned out to yield considerable profits to AT&T. To give you a flavor of the spirit of free inquiry at Bell Labs during that period, I will tell you about the contribution that I am most proud of (which was supported for several years even though it had no monetary value). And I will also mention the contribution that is most often cited of all my papers (which was in collaboration with two mathematicians, and had nothing at all to do with acoustics).

2:45

2pAAd6. Speech and hearing acoustics at Bell Labs. Jont Allen (Univ. of Illinois, Urbana, IL, jba@auditorymodels.org)

A. G. Bell's interest in basic research of speech and hearing was one of the keys to the Bell Lab culture. When the first network circuits were built, speech quality was very low. Research was needed on speech articulation (the probability correct for nonsense speech sounds). George Campbell, a mathematician and ultimate engineer, and expert on Heaviside, extended work of Lord Rayleigh. In 1910 Campbell was the first to generate consonant identification confusion matrices, and show sound grouping (features). Crandall took up this work and attempted (but failed) to define the articulation density over frequency. By 1921 Fletcher had solved Crandall's problem, with the the Articulation Index theory, based on the idea of independent feature perception, across frequency and time. In 1929 he wrote his first book, *Speech and Hearing*, which sold over 5000 copies. His second book, *Speech and Hearing in Communications*, was first released in 1953, after his retirement. Other key people that worked closely with Fletcher were J. C. Steinberg, Munson, French, Galt, Hartley, Kingsbury, Nyquist, Sivian, White, and Wegel. I will try to introduce each of these people and describe their contributions to the speech and hearing field.

3:05–3:20 Break

3:20

2pAAd7. Concert hall acoustics. Manfred Schroeder (Drittes Physikalisches Institut, Univ. Goettingen, Goettingen, Germany)

I will review some work at Bell Laboratories on artificial reverberation and concert hall acoustics including Philharmonic Hall (Lincoln Center for the Performing Arts, New York). I will also touch on sound diffusion by number-theoretic surfaces and the measurement of reverberation time using the music as played in the hall as a ''test'' signal.

3:40

2pAAd8. Transducer research at Bell Labs until 1975. Gerhard M. Sessler (Univ. of Technol., Merckstrasse 25, 64283 Darmstadt, Germany)

Transducer work in the Acoustics Research Department of Bell Labs from about 1960 to 1975 was focussed on electret-based devices. The first such systems were electret headphones which demonstrated the usefulness of this principle in the sense that no dc-bias was required and that the systems were very sensitive, of high quality, and of simple design. This was followed by the invention of the foil-electret microphone in 1962. Initially, these transducers had polyester electrets which, after a thermal charging procedure, kept their charge only for periods of the order of several months. A search for more stable electrets resulted in 1964 in the discovery of fluoroethylenepropylene as a superior material with time constants of the charge decay of hundreds of years under normal environmental conditions. This electret material is still in use in today's transducers. Subsequent activities were directed toward a better understanding of the solid-state properties of these materials. This included, among others, studies of thermally stimulated processes, irradiation effects, and new charging techniques. Later work on transducers centered on the design and characterization of directional microphones, microphone arrays, ultrasonic transducers, touch pads, and other devices useful in telecommunications.

4:00

2pAAd9. Transducer research at Bell Labs 1975 and beyond. James E. West and Ilene Busch-Vishniac (Johns Hopkins Univ., Baltimore, MD, jimwest@jhu.edu)

Condenser microphones, invented at Bell Labs by E. C. Wente and reported in the literature in 1917, are well known for their flat frequency and phase response over a broad range, which make them useful whenever high-quality sound measurements or recordings are required. Two drawbacks, the required high-dc bias and the cost of construction, prevent them from being applied in commercial products such as communication devices, toys, hearing aids, and sound-recording devices. The electret condenser microphone (ECM) eliminates both of the above drawbacks by replacing the needed dc bias with a thin polymer film that is given a permanent charge, the electrical analog of a permanent magnet. Cost of construction is reduced mainly because there is no need to guard against high-voltage breakdown. This technology enables the construction of one- and two-dimensional arrays and up to third-order differential systems for both near- and far-field applications. ECMs are also pliable to most any shape or size, so the construction of continuous strip arrays, transducers for cardiovascular monitoring, MEMS devices, and touch sensors becomes practical. We will discuss some of the experimental systems developed at Bell Labs.

4:20

2pAAd10. Multichannel signal processing at Bell Labs Acoustics Research—Sampled by a postdoc. Walter Kellermann (Multimedia Commun. and Signal Processing, Univ. of Erlangen–Nuremberg, Cauerstr. 7, 91058 Erlangen, Germany)

In the mid 1980's, the first large microphone arrays for audio capture were designed and realized by Jim Flanagan and Gary Elko. After the author joined Bell Labs in 1989, the first real-time digital beamformer for teleconferencing applications was implemented and formed a starting point for the development of several novel beamforming techniques. In parallel, multichannel loudspeaker systems were already investigated and research on acoustic echo cancellation, small-aperture directional microphones, and sensor technology complemented the research scenario aiming at seamless hands-free acoustic communication. Arrays of many sensors and loudspeakers for sampling the spatial domain combined with advanced signal processing sparked new concepts that are still fueling ongoing research around the world—including the author's research group. Here, robust adaptive beamforming has found its way from large-scale arrays into many applications using smaller apertures. Blind source separation algorithms allow for effective spatial filtering without *a priori* information on source positions. Full-duplex communication using multiple channels for both reproduction and recording is enabled by multichannel acoustic echo cancellation combined with beamforming. Recently, wave domain adaptive filtering, a new concept for handling many sensors and many loudspeakers, has been verified for arrays that may well remind some observers of former Bell Labs projects.

4:40

2pAAd11. Room acoustic impulse responses estimation: Challenges and opportunities. Jacob Benesty (Universite du Quebec, INRS-EMT, 800 de la Gauchetiere, Ste. 6900, Montreal, QC H5A 1K6, Canada, benesty@inrs-emt.uquebec.ca), Yiteng (Arden) Huang and Jingdong Chen (Bell Labs., Murray Hill, NJ 07974)

The estimation of room acoustic impulse responses in real-time plays a major role in many important applications such as dereverberation, echo and feedback cancellation, noise reduction, etc. This task is very challenging. In this talk, we will discuss different scenarios of estimation, depending on the application. We will see that when this estimation is reliable, the related applications can give good performances. However, in many applications it is very difficult to have a good estimate of the impulse responses. As a result, performances are usually very poor even with very clever tricks. We will see that opportunities are endless if good progress is made in this important area.

Session 2pAB

Animal Bioacoustics: Animal Bioacoustics General Topics

Robert Kull, Chair
5800 Lake Wright Drive, Suite 101, Norfolk, Virginia 23502

Contributed Papers

1:00

2pAB1. Generalized perceptual features for animal vocalization classification. Patrick J. Clemins and Michael T. Johnson (Marquette Univ., P.O. Box 1881, Milwaukee, WI 53201-1881)

Two sets of generalized, perceptual-based features are investigated for use in classifying animal vocalizations. Since many species, especially mammals, share similar physical sound perception mechanisms which vary in size, two features sets commonly used in human speech processing, mel-frequency cepstral coefficients (MFCCs) and perceptual linear prediction (PLP) analysis, are modified for use in other species. One modification made to the feature extraction process is incorporating the frequency range of hearing and length of the basilar membrane of the animal in order to correctly determine the width and location of the critical band filters used for signal processing. Experimentally determined critical bands (equivalent rectangular bandwidth) and equal loudness curves (audiograms) can also be incorporated directly into the feature extraction process. Experiments are performed on African elephant (*Loxodonta africana*) vocalizations using a hidden Markov model (HMM) based classifier showing increased classification accuracy when using features sets based on the specific animals perceptual abilities compared to the original human perception-based feature sets.

1:15

2pAB2. Acoustic and behavioral analyses of gecker distress vocalizations in young rhesus macaques (*Macaca mulatta*). Erik R. Patel and Michael J. Owren (Psych. of Voice and Sound Res. Lab., Dept. of Psych., Cornell Univ., 211 Uris Hall, Ithaca, NY 14853, erp8@cornell.edu)

Loud, pulsed gecker calls have long been known as one of the most common distress vocalizations produced by young rhesus macaques, but have not been systematically investigated. We therefore examined the rates, acoustics, and behavioral contexts associated with geckering, based on audio-recording and focal-animal observations conducted over 4 years at the California National Primate Research Center. The sample analyzed came from 14 young rhesus macaques, with 74 recorded gecker bouts and 556 total pulses. Callers ranged from 1 to 41 months, who in their first year showed most geckering in months 0–2 (26%) and 2–4 (35%). Females geckered longer and at higher rates than males, as well as showing acoustic differences likely to reflect greater vocal effort (e.g., more doubled and tripled pulses, greater noisiness, and higher first frequency peaks). Geckers were associated with 22 differentiable behavioral contexts, typically occurring when with the mother but without obvious antecedents (31.1%), and when the mother broke physical contact with the youngster (13.5%). The most frequent postcedent context of geckering was that most offspring gained the mother's attention and/or renewed proximity (49%). In combination, the outcomes suggest that geckers primarily function to draw listener attention, and are used by youngsters to elicit maternal response.

1:30

2pAB3. Artiodactyl and Perissodactyl acoustics: Identifying distress calls by farm animals. David G. Browning (Dept. of Phys., Univ. of Rhode Island, Kingston, RI 02881, decibeldb@aol.com) and Peter M. Scheifele (Univ. of Connecticut, Storrs, CT 06269)

There is growing concern for the welfare of farm animals. Vocal signals are discernable in a herd, generally carry over relatively long ranges, and, as Jahns has shown, can be easily automatically detected. Analysis of vocalizations from the two principal farm animal families show, however, that only a few, a pig's squeal, for example, meet Morton's classic criteria for distress. In general, Artiodactyls (cows, sheep, goats, etc.) have tonal bellows or bleats where apparently one vocalization fits many emotional situations. Duration and repetition, as Grandin has suggested, may be the important criteria in indicating stress. In contrast, Perissodactyls vary frequency during some vocalizations, such as a horse whinny, but no direct connection between frequency change and stress has yet been determined. The apparent reliance of Perissodactyles (with keen eyesight) on visual detection of body language appears to limit to some degree the amount of vocalization.

1:45

2pAB4. Vocalization of *Equus asinus*: The hees and haws of donkey brays. David G. Browning (Dept. of Phys., Univ. of Rhode Island, Kingston, RI 02881, decibeldb@aol.com) and Peter M. Scheifele (Univ. of Connecticut, Storrs, CT 06269)

An understanding of the vocalizations of farm animals may provide an important key in the identification of animals under stress. Donkey vocalizations (brays) appear to be unique among the *Equidae* in that significant sound is produced during both air intake (the hee) and air outflow (the haw). Typically these vocalizations, primarily by males, consist of a series of brays, seemingly mechanically produced with little variation, terminating when the animal becomes short of breath—literally a burst of sound. The acoustic character, duration, and sequence (some hee-haw, while others haw-hee) are unique to each animal. As with other Perissodactyles (horses, zebras), and in contrast to the Artiodactyls (cows, sheep, and goats), there can be a significant frequency variation during a vocalization segment (particularly in the ''haw'' phase). It has not been determined what control (if any) the animal has over this. Further study is being made of possible braying variations for a given animal.

2:00

2pAB5. Temporal bone anatomy in *Panthera tigris*. Edward J. Walsh (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131), Darlene R. Ketten, Julie Arruda (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Douglas L. Armstrong, Thomas Curro, Lee G. Simmons (Omaha Henry Doorly Zoo, Omaha, NE 68107), Lily M. Wang (Univ. of Nebraska, Omaha, NE 68182), and JoAnn McGee (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

Preliminary findings suggest that members of *Panthera tigris* subspecies may rely on low-frequency acoustic cues when communicating with conspecifics either in the field or in captivity. This view is supported by the observation that individuals are sensitive to tone bursts in the 300–

500-Hz range and produce significant acoustic energy in an overlapping frequency band in the case of close encounter roars. Other utterances within the vocal repertoire of tigers also contain, and are often dominated by, low frequency acoustic energy that can extend into the infrasonic range. Efforts to determine temporal bone correlates of *P. tigris* bioacoustical features were recently initiated using computerized tomography to assess key aspects of middle and inner ear morphology from a small set of adult Siberian tigers (*P. tigris altaica*) and one neonate. Obvious peripheral auditory specializations were not observed and structures comprising the auditory periphery were consistent with the anatomical character of felids generally. Although cochlear dimensions appeared to be adultlike, or nearly so, in the case of the neonate, other temporal bone features were grossly immature. The relationship between acoustic sensitivity, the spectral character of a subset of close encounter calls and cochlear dimensions will be considered.

2:15

2pAB6. Sound detection and production in the American lobster, *Homarus americanus*: Sensitivity range and behavioral implications. Heidi J. Pye and Winsor H. Watson III (Univ. of New Hampshire, 46 College Rd., Durham, NH 03824)

Many crustaceans including spiny lobsters, some crabs, and a few shrimp are known to produce sounds for a variety of purposes. One brief report and several preliminary studies indicate that American lobsters also produce sounds, and may be capable of detecting acoustic signals. The focus of this study was to (1) quantify the frequency range over which lobsters are capable of detecting sounds and (2) characterize the sounds that lobsters produce. Twelve sexually immature and 11 mature lobsters were tested for their ability to detect frequencies in the range of 20–10 000 Hz. Immature lobsters of both sexes detected sounds in the range of 20–1000 Hz ($>50\%$), while sexually mature lobsters exhibited two distinct peaks in their acoustic sensitivity (20–300 Hz and 1000–5000 Hz). Lobsters of both sexes produced a buzzing vibration when grasped but larger lobsters (120–149 mm in carapace length) vibrated most consistently ($>35\%$ of surveyed lobsters). The greater tendency for sound production in large lobsters may indicate a role in mating behavior. Currently, we are characterizing the acoustical properties of produced sounds and investigating the possibility that the American lobster may produce sounds for more than one purpose. [Work supported by University of New Hampshire Center for Marine Biology.]

2:30

2pAB7. Probability density functions for hyperbolic and isodiachronic locations. John L. Spiesberger (Dept. of Earth and Environ. Sci., Univ. Pennsylvania, 240 S. 33rd St., Philadelphia, PA 19104, johnsr@sas.upenn.edu) and Magnus Wahlberg (Aarhus Univ., DK-8000 Aarhus C, Denmark)

Animal locations are sometimes estimated with hyperbolic techniques by estimating the difference in distances of their sounds between pairs of receivers. Each pair specifies the animal's location to a hyperboloid because the speed of sound is assumed to be spatially homogeneous. Sufficient numbers of intersecting hyperboloids specify the location. A nonlinear method is developed for computing probability density functions for location. The method incorporates *a priori* probability density functions for the receiver locations, the speed of sound, winds, and the errors in the differences in travel time. The traditional linear approximation method overestimates bounds for probability density functions by one or two orders of magnitude compared with the more accurate non-linear method. The non-linear method incorporates a generalization of hyperbolic methods because the average speed of sound is allowed to vary between different receivers and the source. The resulting ''isodiachronic'' surface is the locus of points on which the difference in travel time is constant. Isodiachronic locations yield correct location errors in situations where hyperbolic methods yield incorrect results, particularly when the speed of propagation varies significantly between a source and different receivers.

2:45–3:00 Break

3:00

2pAB8. Information entropy analysis of leopard seal vocalization bouts. John R. Buck (ECE Dept. and SMAST, UMass Dartmouth, 285 Old Westport Rd., North Dartmouth, MA 02747), Tracey L. Rogers (Zoological Parks Board of NSW, Mosman, NSW 2088, Australia), and Douglas H. Cato (Defence Sci. and Technol. Organisation, Pyrmont, NSW 2009, Australia)

Leopard seals (*Hydrurga leptonyx*) are solitary pinnipeds who are vocally active during their brief breeding season. The seals produce vocal bouts consisting of a sequence of distinct sounds, with an average length of roughly ten sounds. The sequential structure of the bouts is thought to be individually distinctive. Bouts recorded from five leopard seals during 1992–1994 were analyzed using information theory. The first-order Markov model entropy estimates were substantially smaller than the independent, identically distributed model entropy estimates for all five seals, indicative of constraints on the sequential structure of each seal's bouts. Each bout in the data set was classified using maximum-likelihood estimates from the first-order Markov model for each seal. This technique correctly classified 85% of the bouts, comparable to results in Rogers and Cato [Behaviour (2002)]. The relative entropies between the Markov models were found to be infinite in 18/20 possible cross-comparisons, indicating there is no probability of misclassifying the bouts in these 18 comparisons in the limit of long data sequences. One seal has sufficient data to compare a nonparametric entropy estimate with the Markov entropy estimate, finding only a small difference. This suggests that the first-order Markov model captures almost all the sequential structure in this seal's bouts.

3:15

2pAB9. Localization of airborne pure tones by pinnipeds. Marla M. Holt, Ronald J. Schusterman, Brandon L. Southall, and David Kastak (Univ. of California, Santa Cruz Long Marine Lab., 100 Shaffer Rd., Santa Cruz, CA 95060, iris@ucsc.edu)

Although all pinnipeds communicate acoustically in air, most previous research on sound localization has been done under water. We have recently shown that several pinniped species localize aerial broadband signals as well as some terrestrial carnivores [Holt *et al.*, J. Acoust. Soc. Am. **113** (2003)]. However, it is unclear which frequencies are particularly important for localization in these animals. In this study, we tested a harbor seal (*Phoca vitulina*) and a California sea lion (*Zalophus californianus*) in a hemianechoic chamber at frequencies ranging between 0.8 and 20 kHz. A left/right procedure was used to measure minimum audible angles (MAAs) corresponding to 75%-correct discrimination. MAAs ranged from approximately 4 to 13 deg in both subjects, with the largest MAAs or poorest acuity measured at the intermediate frequencies tested. These results are consistent with the duplex theory of sound localization in that low-frequency sounds appear to be localized on the basis of interaural time differences, while high-frequency sounds appear to be localized on the basis of interaural intensity differences. Testing with a northern elephant seal (*Mirounga angustirostris*) will provide further insight on the use of binaural cues and head-size effects with respect to localization in pinnipeds.

3:30

2pAB10. Determination of West Indian manatee vocalization levels and rate. Richard Phillips, Christopher Niezrecki, and Diedrich Beusse (Univ. of Florida, Gainesville, FL 32611-6250, niezreck@ufl.edu)

The West Indian manatee (*Trichechus manatus latirostris*) has become endangered partly because of a growing number of collisions with boats. A system to warn boaters of the presence of manatees, based upon the vocalizations of manatees, could potentially reduce these boat collisions. The feasibility of this warning system would depend mainly upon two factors:

the rate at which manatees vocalize and the distance in which the manatees can be detected. The research presented in this paper verifies that the average vocalization rate of the West Indian manatee is approximately one to two times per 5-min period. Several different manatee vocalization recordings were broadcast to the manatees and their response was observed. It was found that during the broadcast periods, the vocalization rates for the manatees increased substantially when compared with the average vocalization rates during nonbroadcast periods. An array of four hydrophones was used while recording the manatees. This allowed for position estimation techniques to be used to determine the location of the vocalizing manatee. Knowing the position of the manatee, the source level was determined and it was found that the mean source level of the manatee vocalizations is approximately 112 dB (*re*:1 Pa) @ 1 m.

3:45

2pAB11. Measurement of Lombard-like response in the beluga whale. Peter M. Scheifele (Animal Sci. Dept., Univ. of Connecticut, 3636 Horsebarn Hill Rd. Ext., Unit 4040, Storrs, CT 06269-4040, scheifel@uconn.edu)

Noise pollution has become recognized as a potential danger to marine mammals in general, and to the St. Lawrence beluga (Delphinapterus leucas) in particular. One method to determine whether noise is having an effect on an animals auditory ability is to observe a natural and repeatable response of the auditory and vocal systems to varying noise levels. This can be accomplished by observing changes in animal vocalizations in response to auditory feedback. A response such as this observed in humans and some animals is known as the Lombard vocal response, which represents a reaction of the auditory system directly manifested by changes in vocalization level. This response is known in humans, songbirds, and some primates. In this research a population of belugas in the St. Lawrence River Estuary was tested to determine whether a vocalization-as-a-function-of-noise phenomenon existed by using hidden Markhov classified vocalizations as targets for acoustical analyses. Correlation and regression analyses of signals and noise indicated that the phenomenon does exist and results of a human subjects experiment along with results from other animal species known to exhibit the response strongly implicate the Lombard vocal response in the St. Lawrence population of beluga.

4:00

2pAB12. Comparative sensitivity analysis of transmission loss in beaked whale environments. Eryn M. Wezensky, James H. Miller, and Robert C. Tyce (Dept. of Ocean Eng., Univ. of Rhode Island, Narragansett Bay Campus, Narragansett, RI 02882)

Scientific literature states that anthropogenic sound, such as mid-frequency sonar, may cause a behavioral response in marine mammals. The degree of response is highly variable and dependent upon many factors, including how sound transmission is influenced by environmental features. The physical parameters of the ocean medium, such as sound speed profile and bathymetry, are important controls of underwater acoustic propagation. Determining the acoustic propagation loss of the ocean environment is an application used to identify and correlate influential environmental factors. This study investigates the sensitivity of acoustic propagation loss based on specific physical characteristics found in five different sites representing beaked whale environments. These sites were chosen with regards to existing data on beaked whale distribution, historical mass stranding records, and presence of mid-frequency sonar activity. A range-independent, ray-tracing acoustic propagation model was used to generate a two-dimensional sound field over a range of 30 km. From the results of this experiment, the acoustic importance of bathymetry and sound speed profile of the five beaked whale environments were identified. Preliminary results from the experimental study will be presented.

4:15

2pAB13. Detecting sperm whale clicks in the presence of ambient and shipping noise using higher order moments. James P. Larue (AFRL/IFEC, Rome, NY 13441), George E. Ioup, and Juliette W. Ioup (Univ. of New Orleans, New Orleans, LA 70148)

The single-receiver detection of various sources using higher order moments is demonstrated. The data for this study come from the Littoral Acoustic Demonstration Center experiment in the northern Gulf of Mexico in the summer of 2001. Results show that in a time-varying environment it may be more meaningful to use non-normalized statistics (e.g., using the fourth central moment rather than the kurtosis). A whale click is detected over a 35-ms window, i.e., with a sampling rate of 11 718 samples per second, 40 samples are used to produce a statistic. In the past, energy and related detectors have been used for the single receiver detection of broadband signals produced by whale clicks. The presence of loud shipping noise, which is also broadband, tends to mask the clicks and make their detection difficult. Using short-time statistics may help mitigate this problem. Analysis with fractal exponents will be shown as well. [Research supported by ONR.]

Session 2pBB

Biomedical Ultrasound/Bioresponse to Vibration and Physical Acoustics: High Intensity Focused Ultrasound II

Cheri Deng, Cochair
Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106

Shahram Vaezy, Cochair
University of Washington, HSB Box 356540, Seattle, Washington 98195-6540

Chair's Introduction—1:00

Invited Papers

1:05

2pBB1. Acoustic hemostasis. Lawrence Crum, Marilee Andrew, Michael Bailey, Kirk Beach, Peter Kaczkowski, Roy Martin, Shahram Vaezy (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105), and Vera Khokhlova (Moscow State Univ., Moscow 119992, Russia)

The application of HIFU to tissue can result in rapid temperature elevations, so much so that protein denaturization can occur. Not only can these elevated temperatures and denatured proteins induce coagulative necrosis, but they can also result in the cessation of blood supply to a particular region. Indeed, when applied to a site of active bleeding, hemostasis can be induced. A group of us at the University of Washington have considered the application of HIFU to active bleeding and have developed methods and devices that offer considerable promise for use in clinical medicine. A brief review of our work in this area will be presented, including our most recent results. [Work supported in part by the U.S. Army and the NIH.]

1:25

2pBB2. Monitoring of HIFU-induced lesions using active imaging methods. Hui Yao and Emad Ebbini (Dept. of Elec. and Computer Eng., Univ. of Minnesota, Minneapolis, MN 55455)

The use of active imaging methods in monitoring the properties of HIFU-induced lesions in freshly excised tissue is investigated. A modern ultrasound scanner is used to collect high-frame rate ultrasound images in conjunction with short-duration localized tissue modification using an external focused transducer. Two modes of reversible tissue modification are used: mechanically pushing the tissue using millisecond pulses and incrementally heating the tissue using 0.5–1 s duration pulses. Speckle tracking techniques are used to estimate the tissue displacements due to shear wave generation (in the case of ms pulse excitation) and temperature change (in the case of 1 s pulse excitation). Experimental results have shown that 20%– 25% change in the shear elastic modulus occurs after lesion formation and appears to be persistent up to 20 min after lesion formation. Absorption appears to increase significantly immediately after heating, but drops nearly to prelesion levels 5 min after heating. A description of the real-time ultrasonic monitoring system will be given. The principles of measurement of both thermal and viscoelastic properties are also described. Illustrative experimental examples from tissue data are given and discussed.

1:45

2pBB3. Fast algorithm for nonlinear acoustics and high-intensity focused ultrasound modeling. Francesco P. Curra, Steven G. Kargl, and Lawrence A. Crum (Ctr. for Industrial and Medical Ultrasound, Appl. Phys. Lab., Univ. of Washington, Seattle, WA, fcurra@apl.washington.edu)

The inhomogeneous characteristics of biological media and the nonlinear nature of sound propagation at high-intensity focused ultrasound (HIFU) regimes make accurate modeling of real HIFU applications a challenging task in terms of computational time and resources. A fast, dynamically adaptive time-domain method that drastically reduces these pitfalls is presented for the solution of multidimensional HIFU problems in complex geometries. The model, based on lifted interpolating second-generation wavelets in a collocation approach, consists of the coupled solution of the full-wave nonlinear equation of sound with the bioheat equation for temperature computation. It accounts for nonlinear acoustic propagation, arbitrary frequency power law for attenuation, multiple reflections, and backscattered fields. The characteristic localization of wavelets in both space and wave number domains allows for accurate simulations of strong material inhomogeneities and steep nonlinear processes at a reduced number of collocation points, while the natural multiresolution analysis of wavelets decomposition introduces automatic grid refinement in regions where localized structures are present. Compared to standard finite-difference or spectral schemes on uniform fine grids, this method shows significant savings in computational time and memory requirements proportional with the dimensionality of the problem. [Work supported by U.S. Army Medical Research Acquisition Activity through the University.]

2:05

2pBB4. Interstitial ablation and imaging of soft tissue using miniaturized ultrasound arrays. Inder R. S. Makin, Laura A. Gallagher, T. Douglas Mast, Megan M. Runk, Waseem Faidi (Ethicon Endo-Surgery, 4545 Creek Rd., Cincinnati, OH 45242, imakin@eesus.jnj.com), Peter G. Barthe, and Michael H. Slayton (Guided Therapy Systems, Mesa, AZ 85202)

A potential alternative to extracorporeal, noninvasive HIFU therapy is minimally invasive, interstitial ultrasound ablation that can be performed laparoscopically or percutaneously. Research in this area at Guided Therapy Systems and Ethicon Endo-Surgery has included development of miniaturized ($\sim$3 mm diameter) linear ultrasound arrays capable of high power for bulk tissue ablation as well as broad bandwidth for imaging. An integrated control system allows therapy planning and automated treatment guided by real-time interstitial B-scan imaging. Image quality, challenging because of limited probe dimensions and channel count, is aided by signal processing techniques that improve image definition and contrast. Simulations of ultrasonic heat deposition, bio-heat transfer, and tissue modification provide understanding and guidance for development of treatment strategies. Results from *in vitro* and *in vivo* ablation experiments, together with corresponding simulations, will be described. Using methods of rotational scanning, this approach is shown to be capable of clinically relevant ablation rates and volumes.

2:25

2pBB5. Clinical exposure protocols for HIFU treatments of cancer of the liver and kidney. Gail R. ter Haar (Phys. Dept., Royal Marsden Hospital, Sutton, Surrey SM2 5PT, UK), Ian H. Rivens (Royal Marsden Hospital, Surrey SM2 5PT, UK), James E. Kennedy (Churchill Hospital, Oxford OX3 7LJ, UK), and Feng Wu (Chongqing Univ. of Medical Sci., Chongqing, PRC)

High-intensity focused ultrasound surgery (HIFU) is a technique that is finding increasingly widespread clinical application. The potential of HIFU for the noninvasive treatment of deep-seated tumors is being explored. Treatment of these tumors uses an extracorporeal approach. A number of different devices are now available for delivering clinical HIFU treatments. This paper sets out methods for characterization of HIFU transducers and exposure protocols that would allow treatments to be compared. Clinical trials for the treatment of tumors of the liver and kidney are underway in the UK at the Royal Marsden Hospital, London and the Churchill Hospital, Oxford. This paper describes the physical and technical characteristics of the two devices being used to deliver treatments, and uses these as a basis for discussing the ultrasonic procedures that can be used, and the rationale for their choice. Clinical reference is given for illustrative purposes. The wide variety of exposure protocols that is being used in different clinical centers means that there is at present no uniformity of reporting in the literature. This paper therefore sets out recommended information that should be included in order that transducers and treatments can be compared.

2:45

2pBB6. Prostate cancer treated with HIFU: A 10-year experience. Jean-Yves Chapelon, Laura Curiel, and Albert Gelet (INSERM U556, 151, Cours Albert Thomas, 69424 Lyon, Cedex 03, France, chapelon@lyon.inserm.fr)

Objectives: To evaluate our overall experience in HIFU in the prostate cancer application, for its two main indications: as a primary care and as a salvage therapy after radiation. Material and method: Group 1: patients treated with HIFU as primary care for localized prostate cancer, with a baseline PSA level lower than 30 ng/ml, and with at least 1 year of follow-up were included in this analysis. Group 2: patients with a local recurrence after radiation. They were all treated using the Ablatherm HIFU device (EDAP S.A., France). Results: Group 1: $n = 242$ patients. The mean nadir PSA was 0.63 sd 1.30 ng/ml, and further follow-up evidenced a 81% negative biopsy rate. These results are influenced by the usual prognostic factors (clinical stage, baseline PSA, Gleason score), and, as for the other treatment option, the nadir allows prediction of the patient outcome. Group 2: $n = 71$ patients. After HIFU treatment, 80% of the patients presented negative biopsies, and 61% had a nadir PSA level lower than 0.5 ng/ml, obtained within 3 months. No rectal injury occurred since the use of specific parameters. Conclusion: Transrectal HIFU is a valuable option for prostate cancer, for primary care as well as for salvage therapy.

3:05–3:20 Break

Contributed Papers

3:20

2pBB7. Real-time motion correction combined with electronic beam steering strategies optimized for large tumors: First *in-vitro* HIFU experiments in large moving targets. Mathieu Pernot, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, E.S.P.C.I., 10 rue Vauquelin 75231 Paris Cedex 05, France, mathieu.pernot@loa.espci.fr)

A method for tracking in realtime the 3D motion of tissues is combined with a 2D high-intensity focused ultrasound (HIFU) multichannel system in order to correct for respiratory motion during HIFU therapy. Displacements estimation is based on an accurate 3D ultrasonic speckle tracking technique. A correction is achieved in real time by adjusting the transmit delays of each channel of the HIFU system. *In vitro* HIFU experiments combined with motion correction are performed in fresh biological tissues. The accuracy of the HIFU targeting is clearly improved. Moreover, this technique permits an important reduction of the treatment time for moving tumors. To generate large thermal lesions, the focus is scanned electronically over a large region. Different heating strategies are investigated in moving tissues: the shot-by-shot treatment interleaved with cooling periods, and the continuously scanned treatment of the whole region. The temperature distribution and the thermal dose are computed using a 3D thermal diffusion code combined with a 3D wave diffraction algorithm and compared to experimental results. Using motion correction, the necrosis threshold is achieved 3 times faster than without motion correction. This should lead to an important time reduction for the treatment of abdominal tumors.

3:35

2pBB8. Therapy/imaging array-based system and technology for intense ultrasound surgery. Peter G. Barthe, Michael H. Slayton, Paul M. Jaeger (Guided Therapy Systems, 33 S. Sycamore St., Mesa, AZ 85202, p.barthe@guidedtherapy.com), Inder R. S. Makin (Ethicon Endo-Surgery, Cincinnati, OH 45242), Laura A. Gallagher, T. Douglas Mast, Megan M. Runk, and Waseem Faidi (Ethicon Endo-Surgery, Cincinnati, OH 45242)

Minimally invasive, miniature (2.2- × 50-mm aperture, 3.3-mm diameter) dual-mode linear arrays have been developed into low-cost disposable probes with high acoustic power output (120 W/cm^2 at the source), high transmit efficiency (>65% typical), and good imaging performance (50% fractional bandwidth, >100-mm-deep field of view). These therapy/imaging probes have been integrated into a flexible intense ultrasound surgery platform which also includes conventional diagnostic imaging probes. A system architecture has been developed which includes a 64-channel therapy driver with software selection of array aperture and phasing ($\lambda/16$), frequency (0.5–8 MHz), drive amplitude (5 W/channel, nominal), rotational steering (± 180 deg), and temporal sequencing/switching of imaging/therapy/monitoring modes. System software includes graphical and text-based script mode control of therapeutic treatment. Real-time monitoring of electric power per channel, temperature sensors, and thermal effects provide a range of feedback and safety. Numerous system and probe technological issues such as electrical interconnect and matching, acoustic coupling, thermal control, and maintaining probe efficiency have been addressed. The array-based imaging/therapy system has produced encouraging results in preclinical studies of bulk tissue ablation and imaging.

3:50

2pBB9. Dual-modality phased array systems for hyperthermia cancer therapy. Robert McGough (Dept. of Elec. and Computer Eng., Michigan State Univ., 2120 Eng. Bldg., East Lansing, MI 48824) and Thaddeus Samulski (Duke Univ. Medical Ctr., Durham, NC 27710)

External ultrasound (US) phased arrays, as applied to hyperthermia cancer therapy, are generally local heating devices, whereas external radio-frequency (RF) electromagnetic (EM) phased array applicators are primarily designed for regional heating. Clinical applications of external ultrasound heating devices are presently limited to smaller tumors, and the inclusion of an external RF/EM applicator increases the size of the heated volume. When these two modalities are combined in a hybrid RF/US phased array structure, uniform heating is facilitated in larger tumors. This is demonstrated in computer simulations of thermal therapy in the breast, where adding an external RF/EM component to an ultrasound phased array in a hybrid structure produces temperatures of 43 °C in nearly half of a 9-cm-diam tumor volume. If applied alone, neither modality achieves 43 °C in such a large volume without overheating intervening tissues. This result is demonstrated in two hybrid RF/US applicator prototypes designed for hyperthermia treatments of locally advanced breast cancer. In each example, hybrid RF/US phased arrays achieve higher temperatures in larger tumor volumes by selectively delivering power to different regions in the tumor with each modality.

4:05

2pBB10. A multi-element high intensity focused ultrasound transducer: Design, fabrication, and testing. Shahram Vaezy, Robert Held, Blake Miller (Appl. Phys. Lab., Univ. of Washington, Box 355640, vaezy@apl.washington.edu), and Gerard Fleury (Imasonic Corp., Besancon, France)

The goal of this project is to develop an intra-cavity image-guided high intensity focused ultrasound (HIFU) device using piezocomposite technology and commercially available ultrasound imaging. The HIFU array, manufactured by Imasonic Corporation, is an 11-element annular phased array, with a focal length range of 30–60 mm, and operating frequency of 3 MHz (bandwidth of 1 MHz). The imaging probe (C9-5, Philips) is configured such that the focal axis of the HIFU beam was within the image plane. The array includes six complete central rings and five side-truncated peripheral rings, all with the natural radius of curvature of 50 mm. Impedance of all elements is approximately 50 ohms (10% accuracy for real and imaginary parts). Cross coupling between adjacent elements is less than, −40 dB. High power measurements showed more than 75% efficiency, at surface intensity of 2.66 W/cm^2. Schlieren imaging showed effective focusing at all focal lengths (30–60 mm). The image-guided HIFU device requires water or hydrogel coupling, and possibly water cooling. The results of the full characterization for lesion formation in tissue-mimicking phantoms and biological tissues will be presented. Possible applications include uterine fibroids, abnormal uterine bleeding, and intraoperative hemostasis of occult hemorrhage.

4:20

2pBB11. High-intensity focused ultrasound for liver biopsy hemostasis. Cheri Deng, Hesheng Wang, Yun Zhou (Dept. of Biomed. Eng., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106-7207, cxd54@cwru.edu), Vikram Dogra, Agata Exner, Shweta Bhatt, John Haaga, and Nicholas Stowe (Univ. Hospitals of Cleveland, Cleveland, OH 44106)

In vivo experiments were conducted to demonstrate the feasibility of HIFU application to control postliver biopsy hemorrhage. Yorkshire pigs were anesthetized and their livers were surgically exposed. Core biopsies ($n=74$) were performed on the exposed hepatic parenchyma with 14-gauge ($n=41$) and 18-gauge ($n=33$) core biopsy needles that were inserted 1.5–2 cm deep into the liver. Hemorrhage was determined from the weight of the blood collected from each biopsy puncture site using surgical sponges immediately after biopsy needle retraction. To stop hemorrhage, immediate HIFU was applied to the needle entry site ($n=44$) after needle retraction. HIFU was generated using a piezoelectric (PZT) transducer (diameter=42 mm, F number=1.2) at 4.23 MHz. Whole-blood clotting times were measured at various times throughout the experiments. Mean blood loss from control biopsy sites using a 14-gauge needle ($n=18$) was 1.78 g, while mean blood loss using an 18-gauge needle ($n=10$) was 1.22 g (two 14-gauge-needle control biopsies were excluded). Virtually no blood loss was measured from the biopsy needle entry site after HIFU application for both 14- and 18-gauge-needle biopsies. Ultrasound imaging demonstrated a marked difference between control sites and HIFU-treated sites where successful hemostasis was achieved.

4:35

2pBB12. A model-based noninvasive temperature estimation technique for monitoring HIFU therapy using backscattered ultrasound. Ajay Anand and Peter J. Kaczkowski (Ctr. for Industrial and Med. Ultrasound, APL, Univ. of Washington, 1013 40th St. NE, Seattle, WA 98105, ajaya@apl.washington.edu)

A new noninvasive temperature estimation technique using the bioheat transfer equation (BHTE) to constrain inversion of ultrasonic travel time data has been developed for HIFU therapy monitoring. Initial estimates for BHTE parameters and ultrasonic travel time temperature dependence $\delta\tau(T)$ are noninvasively obtained *in situ* using ultrasound backscatter signals acquired during two probe heating exposures performed prior to therapy: one to obtain medium thermal parameters, the other to estimate the magnitude of the local HIFU heat source. These estimates are used in BHTE simulations and compared to local travel time changes measured periodically during therapy using a medical ultrasound imager. Quantitative 2-D temperature maps are obtained by constrained data fitting. Preliminary results obtained from *in vitro* experiments in tissue mimicking phantoms will be presented. Temperature maps are computed throughout therapy delivery and posttreatment cooling periods, and validated against independent temperature measurements obtained using thermocouples placed close to but not at the HIFU focus. This model-based technique permits temperature estimation throughout the entire therapeutic range and is completely noninvasive, and is thus a departure from previously reported techniques. This ultrasound-based approach could serve as an alternative to MRI-based temperature estimation methods in a clinical setting.

4:50

2pBB13. Dosimetry measurement of HIFU field using a fiber-optical probe hydrophone. Liang Zhai, Yufeng Zhou, and Pei Zhong (Dept. of Mech. Eng. and Mater. Sci., Duke Univ., Box 90300, Durham, NC 27708, pzhong@duke.edu)

Characterization of the acoustic field of high-intensity focused ultrasound (HIFU) transducers by conventional PVDF membrane and needle hydrophones is problematic due to limited bandwidth, spatial averaging, and damage to the hydrophone. Here, we report the use of a self-calibrated fiber-optical probe hydrophone (FOPH-500) for HIFU dosimetry measurement. The hydrophone (0.1-mm sensing element) was scanned in the focal volume of a 1.1-MHz HIFU transducer ($F=63$ mm, f number$=0.9$) at 0.2-mm steps using a computer-controlled 3D positioning system. When the input voltage V_{p-p} applied to the transducer was increased from 28 to 225 volts, the peak compressive and tensile pressure values at the transducer focus were found to be $P^{+}=1.7$–18.9 MPa and $P^{-}=-1.33$–-8.14 MPa, respectively. The corresponding spatial peak intensities were calculated to be $I_{SP}=69$–3968 W/cm^2. Nonlinear propagation with harmonics generation was dominant at high intensity levels, leading to a reduced -6-dB beamwidth of the compressive wave from 1.8 to 1.3 mm and an increased -6-dB beamwidth of the tensile wave from 1.6 to 1.8 mm. Overall, FOPH-500 was found to be a reliable tool for characterizing the acoustic field of HIFU transducers.

TUESDAY AFTERNOON, 25 MAY 2004 LIBERTY 5, 1:30 TO 4:00 P.M.

Session 2pED

Education in Acoustics: Apparatus for Teaching Acoustics, 1929 and Before

Thomas D. Rossing, Cochair
Physics Department, Northern Illinois University, De Kalb, Illinois 60115

Peter L. Hoekje, Cochair
Department of Physics and Astronomy, Baldwin-Wallace College, 275 Eastland Road, Berea, Ohio 44017

Chair's Introduction—1:30

Invited Papers

1:35

2pED1. Apparatus for studying wave motion and sound at the University of Nebraska–Lincoln's "Historical Scientific Instrument Gallery." Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska, Peter Kiewit Inst., Omaha, NE 68182-0681, lwang4@unl.edu) and M. Eugene Rudd (Univ. of Nebraska, Lincoln, NE 68588)

The University of Nebraska–Lincoln's "Historical Scientific Instrument Gallery," compiled by the second author in 1998, contains approximately 700 inventoried items and may be visited on-line at *http://physics.unl.edu/outreach/histinstr/*. Amidst the collection are several acoustical instruments that were used in the early 1900s. These include equipment that demonstrate wave motion (traveling wave machine, mercury ripple dish, vibration microscope), wave interference (interference machine), resonance conditions (Helmholtz resonators, vibrating rods, singing flames, sonometer), and sound generation (Galton's whistles, high-frequency tuning forks, large tuning forks, organ pipes, siren saw). A review of the equipment and the history of their use at the University of Nebraska are discussed. Much of the equipment was superbly manufactured by the Max Kohl/Chemnitz Company in Germany and Rudolph Koenig in France. Pages from the Max Kohl/Chemnitz equipment catalogs of 1910 and 1925 helped to characterize several of the pieces and are shown in this presentation.

1:55

2pED2. Early 20th century acoustics apparatus in Iowa. Roger J. Hanson (Dept. of Phys., Univ. of Northern Iowa, Cedar Falls, IA 50613, roger.hanson@cfu.net)

In the first half of the 20th century G. W. Stewart was a physics faculty member at the University of Iowa (UI) with a distinguished record of research and teaching, especially in acoustics. Much of his research focused on the design and use of several types of acoustical filters. Some apparatus which he developed or utilized are still housed in the Department of Physics and Astronomy or are available in detailed diagrams. Demonstration apparatus (apparently homemade) from his era are still available for use. Carl E. Seashore, a renowned psychologist also at UI in the early 20th century, had interdisciplinary interests linking psychology, speech and hearing, music, and acoustics. He was responsible for obtaining an Henrici harmonic analyzer, a mechanical Fourier analyzer manufactured in Switzerland, a special grant from the state legislature during Depression conditions provided the funding. It resides in the Department of Speech Pathology and Audiology at UI. The Grinnell College Physics Historical Museum houses a set of 18 Helmholtz resonators and a Savart bell and resonator. Apparatus at Iowa State University, the University of Northern Iowa, and other Iowa institutions will also be described. Pictures and diagrams as well as some actual apparatus will be exhibited.

2:15

2pED3. Rudolph Koenig's workshop of sound. David A. Pantalony (Dibner Inst., MIT, 38 Memorial Dr., Cambridge, MA 02139, pantalony@mit.edu)

Rudolph Koenig's workshop was a busy meeting place for instruments, ideas, experiments, demonstrations, craft traditions, and business. Starting around 1860 it was also the place in Paris where people discovered the new science of sound emerging from the studies of Hermann von Helmholtz in Germany. Koenig built Helmholtz's ideas into apparatus, created new instruments, and spread them throughout the scientific and musical world. Through his own research, he also became Helmholtz's strongest critic. This paper looks at the activities of this unique space, and, in particular, how it contributed to the protracted disputes over an elusive acoustical phenomenon called the combination tone. Many of these instruments became standard teaching and demonstration apparatus.

2:35

2pED4. Harmonic analysis utilizing a Phonodeik and an Henrici analyzer. William J. Fickinger (Dept. of Phys., Case Western Reserve Univ., Cleveland, OH 44106, wjf@case.edu), Roger J. Hanson (Univ. of Northern Iowa, Cedar Falls, IA 50614), and Peter L. Hoekje (Baldwin-Wallace College, Berea, OH 44017)

Dayton C. Miller of the Case School of Applied Science assembled a series of instruments for accurate analysis of sound [D. C. Miller, J. Franklin Inst. **182**, 285–322 (1916)]. He created the Phonodeik to display and record sound waveforms of musical instruments, voices, fog horns, and so on. Waveforms were analyzed with the Henrici harmonic analyzer, built in Switzerland by G. Coradi. In this device, the motion of a stylus along the curve to be analyzed causes a series of spheres to rotate; two moveable rollers in contact with the nth sphere record the contributions of the sine(nx) and cosine(nx) components of the wave. Corrections for the measured spectra are calculated from analysis of the response of the Phonodeik. Finally, the original waveform could be reconstructed from the corrected spectral amplitudes and phases by a waveform synthesizer, also built at Case. Videos will be presented that show the motion of the gears, spheres, and dials of a working Henrici analyzer, housed at the Department of Speech Pathology and Audiology at the University of Iowa. Operation of the Henrici analyzer and the waveform synthesizer will be explained.

2:55

2pED5. Acoustic teaching apparatus before 1929 at the Case School of Applied Science. Peter L. Hoekje (Dept. of Phys. and Astron., Baldwin–Wallace College, 275 Eastland Rd., Berea, OH 44017, phoekje@bw.edu) and William Fickinger (Case Western Reserve Univ., Cleveland, OH 44106)

The acoustics apparatus found in the Physics Department of the Case School of Applied Science in the first decades of the 20th century included many items common to other acoustical teaching laboratories, such as organ pipes, tuning forks, Helmholtz resonators, sirens, and manometric flame sound analyzers. The European instrument makers Rudolf Koenig and Max Kohl supplied much of this. Equipment built at Case included the phonodeik, which Dayton C. Miller designed in 1908, and the waveform synthesizer. Miller supplied detailed descriptions of the operations of all this equipment in papers and books. In the phonodeik (to show sound), sound deflects a thin glass diaphragm, which by a silk thread turns a mirror on an axle, causing a spot of light to move across film or a projection screen. A working model of the phonodeik has been reconstructed from pieces of two original ones, and will be demonstrated. Photographs of other extant instruments in the collection, and a selection from Millers lantern slides, will be displayed.

Contributed Papers

3:15

2pED6. A model for precalculus students to determine the resonance frequency of a trumpet mouthpiece. Robert C. Chapman (Jefferson County Public Schools, 1829 Denver West Dr., Golden, CO 80401-3120)

The trumpet mouthpiece as a Helmholtz resonator is used to show precalculus students a mathematical model for determining the approximate resonance frequency of the mouthpiece. The mathematics is limited to algebra and trigonometry. Using a system of mouthpieces that have interchangeable cups and backbores, students are introduced to the acoustics of this resonator. By gathering data on 51 different configurations of mouthpieces, the author modifies the existing Helmholtz resonator equation to account for both cup volumes and backbore configurations. Students then use this model for frequency predictions. Included are how to measure the different physical attributes of a trumpet mouthpiece at minimal cost. This includes methods for measuring cup volume, backbore volume, backbore length, throat area, etc. A portion of this phase is designed for students to become acquainted with some of the vocabulary of acoustics and the physics of sound.

3:30

2pED7. Acoustic engineering at Universidad de las Americas, Ecuador. Luis A. Bravo, Jaime O. Naranjo, and Alberto Tassara (Coln 338 y 6 de Diciembre, Quito, Ecuador, lbravo@uamericas.edu.ec)

Acoustics, like science, an instrument to develop new technologies, comfortable atmospheres, and pleasant sounds, has not had a sufficient push in Ecuador. The shortage of professionals in the area, and the social ignorance of the advances and benefits of acoustics have been part of the problem. The University of the Americas has taken the initiative to develop an undergraduate program—only in the country—of sound and acoustics engineering, to contribute to the formation of professional fu-

tures that fortify the recent labor market in the areas of audio, professional, and acoustic engineering. This work presents/displays the results of the studies made for the creation of the race, the curricular mesh, and its projections.

3:45

2pED8. Comparison of Russian and American educational programs in acoustics. Nikolai Zagrai and Vladislav Zakharevich (Taganrog State Univ. of Radio-Eng., GSP-17A, 44 Nekrasovsky St., Taganrog 347928, Russia, znp@tsure.ru)

Recent progress in finding parallels between Russian and American professional education programs has shown that particular attention should be paid to the differences and similarities in engineering education. The efforts of the Taganrog State University of Radio-Engineering (TSURE) are directed towards accreditation of its engineering programs in ABET. The Russian Bachelor program in acoustics was compared to the ABET-accredited American program in electrical engineering. A direct comparison has shown that Russian program has: (a) more credit hours in the course work under the liberal arts section (factor of 2); (b) traditionally more credit hours in mathematics and fundamental science (factor of 1.5); (c) same number of credit hours for the general engineering classes; (d) less number of credit hours for the specialized engineering classes. However, by examining the actual class contents of the specialized engineering classes in the American program, we observed that some classes contain material which, in Russian program, is usually attributed to the general engineering. Recent revision of the acoustic education program at TSURE aimed to increase its flexibility, support students' mobility, and facilitate ABET accreditation.

TUESDAY AFTERNOON, 25 MAY 2004 CONFERENCE ROOM K, 1:00 TO 3:20 P.M.

Session 2pMU

Musical Acoustics and Signal Processing in Acoustics: Digital Signal Processing Methods for Restoring, Enhancing and Manipulating Music Recordings

James W. Beauchamp, Chair

School of Music, Department of Electrical and Computer Engineering, University of Illinois—Urbana/Champaign, 1114 West Nevada, Urbana, Illinois 61801-3859

Invited Papers

1:00

2pMU1. Retrieving the sources in historical sound recordings. George Brock-Nannestad (Patent Tactics, Resedavej 40, DK-2820 Gentofte, Denmark, pattac@image.dk)

Broadly speaking, historical recordings (ca. 1880–1950) are considered low quality, and when they are reissued it is regarded as an improvement when noises are removed, or according to modern psychoacoustic criteria, pushed into frequency bands where they are masked. However, a recording is a report of an acoustic event, and some of the elements or features of the original sounds must have been available to contemporary listeners, or they would not have accepted the recordings in the first place. Similarly, we may want to retrieve these elements, and for that purpose it is important to identify them. The paper points out some of the fundamental elements of the sounds that have been recorded and discusses the degree to which they may be retrieved in modern replay, either directly from original recordings or via signal-processed transfers. Reference is made to work by D. C. Miller based on his early recognition that the recording process transforms the elements it is desired to retrieve. In a similar manner that Miller compensated his recordings for measurement purposes, we may today compensate early recordings for replay purposes. The presentation will be accompanied by very short comparative excerpts of historical sound recordings.

1:20

2pMU2. An overview of statistical-model-based techniques for audio restoration. Patrick J. Wolfe (Dept. of Eng., Univ. of Cambridge, Trumpington St., Cambridge CB2 1PZ, UK)

This presentation will describe the state of the art in model-based approaches to audio restoration, with an emphasis on results and an explanation of theory suitable for the nonspecialist. An overview of work conducted over the past several years by members of the Cambridge Signal Processing Group will be presented, including techniques for the detection, correction, and interpolation of degraded or missing audio data. The latest advances in a fully Bayesian statistical methodology for modeling of music and speech signals will also be detailed, including a common unifying framework for the treatment of both global and localized audio degradations.

1:40

2pMU3. A two-stage approach to removing noise from recorded music. Jonathan Berger, Maxim J. Goldberg, Ronald C. Coifman (CCRMA, Stanford Univ., Stanford, CA 94305, brg@ccrma.stanford.edu), Maxim J. Goldberg (Ramapo College of New Jersey, Mahwah, NJ 07430), and Ronald C. Coifman (Yale Univ., New Haven, CT 06520)

A two-stage algorithm for removing noise from recorded music signals (first proposed in Berger *et al.*, ICMC, 1995) is described and updated. The first stage selects the ''best'' local trigonometric basis for the signal and models noise as the part having high entropy [see Berger *et al.*, J. Audio Eng. Soc. **42**(10), 808–818 (1994)]. In the second stage, the original source and the model of the noise

obtained from the first stage are expanded into dyadic trees of smooth local sine bases. The best basis for the source signal is extracted using a relative entropy function (the Kullback–Leibler distance) to compare the sum of the costs of the children nodes to the cost of their parent node; energies of the noise in corresponding nodes of the model noise tree are used as weights. The talk will include audio examples of various stages of the method and proposals for further research.

2:00

2pMU4. Reconstruction of mechanically recorded sound by image processing. Vitaliy Fedeyev, Carl Haber, Zachary Radding (Lawrence Berkeley Natl. Lab., Berkeley, CA 94270), Christian Maul (TaiCaan Technologies Ltd., Southampton, UK), John McBride (Univ. of Southampton, Southampton, UK), and Mitchell Golden

Audio information stored in the undulations of grooves in a medium such as a phonograph record or cylinder may be reconstructed, without contact, by measuring the groove shape using precision optical metrology methods and digital image processing. In this approach, audio signal processing is accomplished by two- or three-dimensional image analysis and processing. The viability of these methods was recently demonstrated on a 78 rpm shellac disc using two-dimensional image acquisition and analysis methods [V. Fadeyev and C. Haber, J. Audio Eng. Soc. **51**(12), 1172–1185 (2003)]. The present work expands on these results. A three-dimensional reconstruction of mechanically recorded sound is reported. The source material, an Edison cylinder, was scanned using confocal microscopy and resulted in a faithful playback of the recorded information. Methods to accelerate the scan rates and make these techniques practical for use in working archives are reported as well. [Work supported by the Laboratory Technology Research Program (SC-32), within the Office of Science, U.S. Department of Energy under Contract No. DE-AC03-76SF00098.]

2:20

2pMU5. Source identification and manipulation in stereo music recordings using frequency-domain signal processing. Carlos Avendano (Creative Adv. Technol. Ctr., Scotts Valley, CA 95066)

A short-time frequency domain framework for source identification, separation, and manipulation in stereo music recordings is presented. Using a simplified model of the stereo mix, a similarity measure between the short-time fourier transforms (STFTs) of the input signals is computed to identify time-frequency regions occupied by each source based on the panning coefficients assigned to it during the mix. Individual sources are identified and manipulated by clustering time-frequency components with a given panning coefficient and frequency range. After modification, an inverse STFT is used to synthesize a time-domain processed signal. Applications of the technique to source suppression, enhancement and repanning will be described, and audio demonstrations will be presented to illustrate the results.

2:40

2pMU6. SmartMusicKIOSK: Music-playback interface based on chorus-section detection method. Masataka Goto (Natl. Inst. of Adv. Industrial Sci. and Technol. (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan, m.goto@aist.go.jp)

This paper describes *SmartMusicKIOSK*, a new music-playback interface for trial listening. Traditionally in music stores, customers often search out the chorus or ''hook'' of a song by repeatedly pressing the fast-forward button, rather than passively listening to the music. This activity is not well-supported by current technology. This research achieves a function for jumping to the chorus section and other key parts of a song, plus a function for visualizing song structure. These functions eliminate the hassle of searching for the chorus and make it easier for a listener to find desired parts of a song, thereby facilitating an active listening experience. This interface, which enables a listener to look for a section of interest by interactively changing the playback position, is useful not only for trial listening but also for more general purposes in selecting and using music. The proposed functions are achieved through an automatic audio-based chorus-section detection method that can detect all the chorus sections by analyzing relationships between various repeated sections in a song. It can also detect modulated chorus sections by introducing an acoustic similarity that enables modulated repetition to be judged correctly. The results of implementing this method in SmartMusicKIOSK have demonstrated its usefulness.

3:00

2pMU7. Music processing above and below the fundamental frequency. Marios Athineos and Daniel P. W. Ellis (LabROSA, Dept. of Elec. Eng., Columbia Univ., New York, NY 10027)

Because of the centrality of pitch to music, most music signal processing has been focused on pitch-related timescales, e.g., short-time analysis over 30–50-ms windows. But there is, of course, important information at other scales as well. By looking at ways to model and extract very fine time-scale information one can extract attributes that lead to the perception of ''texture'' in unpitched instruments such as maracas. At the other extreme, interesting effects can be revealed by investigating the effect of ''modulation-domain'' processing, such as filtering subband energy envelopes in the 0.1–10-Hz range. A single representation, frequency-domain linear prediction, gives convenient access to both these scales: its use in music analysis/synthesis will be illustrated by examples.

Session 2pNSa

Noise and Psychological and Physiological Acoustics: Noise Impact Evaluation: Old and New II

Brigitte Schulte-Fortkamp, Cochair
Institute of Technical Acoustics, Technical University Berlin, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany

Klaus Genuit, Cochair
HEAD Acoustics GmbH, Eberstrasse 30a, Herzogenrath 52134, Germany

Chair's Introduction—1:15

Invited Papers

1:20

2pNSa1. Meaningless artificial sound and its application in urban soundscape research. Bert De Coensel and Dick Botteldooren (Acoust. Group, Dept. of Information Technol., Ghent Univ., Belgium, bert.decoensel@intec.ugent.be)

Urban areas are increasingly being overwhelmed with uninteresting (traffic) noise. Designing a more matching soundscape for urban parks, quiet backyards, shopping areas, etc., clearly deserves more attention. Urban planners, being architects rather than musical composers, like to have a set of "objective" indicators of the urban soundscape at their disposal. In deriving such indicators, one can assume that the soundscape is appreciated as a conglomerate of sound events, recognized as originating from individual sources by people evaluating it. A more recent line of research assumes that the soundscape as a whole evokes particular emotions. In this research project we follow the latter, more holistic view. Given this choice, the challenge is to create a test setup where subjects are not tempted to react to a sound in a cognitive way, analyzing it to its individual components. Meaningless sound is therefore preferred. After selection of appealing sounds for a given context by subjects, objective indicators can then be extracted. To generate long, complex, but meaningless sound fragments not containing repetition, based on a limited number of parameters, swarm technology is used. This technique has previously been used for creating artificial music and has proved to be very useful.

1:40

2pNSa2. Subjective soundscapes qualitative research in the experience and evaluation of environmental noise. Uwe Flick (Alice Salomon Univ. of Appl. Sci., Alice-Salomon-Platz 5D-12627, Berlin, Germany, Uwe.Flick@Tu-Berlin.de)

If the subjective experience and evaluation of environmental noise shall be considered and integrated into the current soundscape research, the use of qualitative research methods used in sociology and psychology will become necessary. A triangulation of research methods for measuring objective noise and for the subjective evaluation of noises and sounds on the background of subjective meanings of health and healthy living will be a fruitful way to a more comprehensive understanding of the phenomenon of soundscapes in the context of health and quality of life. In this contribution, a selection of qualitative research methods will be presented that allows for analyzing subjective experiences with environmental noise. Interviews focusing on narratives of episodes and situations (e.g., the episodic interview, Flick, 2002) will be outlined. Issues of how to assess the quality of qualitative research and its results will be addressed and finally the benefits and limits of the triangulation of different methods (e.g., interviews and focus groups or interviews and physical measures) will be discussed. Research experiences from the author's recent studies on health concepts of health professionals will be used for illustration.

2:00

2pNSa3. A cognitive approach to soundscape research. Daniele Dubois, Catherine Guastavino, Valerie Maffiolo, Manon Raimbault (CNRS LCPE/LAM 11 rue de Lourmel (F) 75015 Paris, France, ddubois@ccr.jussieu.fr), Catherine Guastavino (McGill Univ., Canada), Valerie Maffiolo (FTR&T Lannion, France), and Manon Raimbault (INRETS, Lyon, France)

The present research on cognitive categories for soundscapes focuses on their interpretations and can be seen as mediating between individual sensory experiences and collective representations shared in language and elaborated as knowledge. Results of field inquiries in Paris, Lyon, and Nantes are presented together with results from categorization of recorded soundscapes in laboratory conditions. Categories were identified by means of linguistic analyses of verbal comments and mathematical analyses of similarity judgments. Results indicate that people categorize environmental sounds on the basis of semantic features, namely source identity and pleasantness judgments, rather than perceptual features. Effects of noise on human subjectivity cannot be quantitatively measured thoroughly in terms of physical parameters: auditory judgments depend upon the meaning attributed to acoustic phenomena and noise sources, rather than on inherent properties of the acoustic signal. These findings highlight the fact that an acoustic phenomenon can be diversely conceptualized and lexicalized as cognitive representations. Finally, methodological and theoretical consequences of these findings are established as the basis for further research on soundscape, in order to account not only for noise annoyance but also for sound quality of urban life.

2:20

2pNSa4. Soundscapes and their influence on inhabitants—New findings with the help of a grounded theory approach. Andre Fiebig and Brigitte Schulte-Fortkamp (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany, a.fiebig@gmx.de)

Recent research work has shown the relevance of soundscapes with respect to community noise. A study was conducted aiming at clarifying and improving the comprehension between the acoustical stimulus and assessment, response, attitude, and perception of inhabitants in urban environments regarding their urban soundscapes. Narrative and issue-centered interviews were carried out to identify common perceptive patterns. Therefore subjects were required to articulate their opinions and sensations about their surroundings including environmental sound. At the same time acoustic measurements were performed. Therefore the study employs methodology of qualitative-interpretative research, since perception is constantly embedded in social contexts. By means of the ''grounded theory'' approach, a general methodology of qualitative social analysis, individual and interindividual determinants have been explored in the process of evaluating urban soundscapes based on a new integrative diagram. The new model will be presented.

2:40

2pNSa5. The acoustical diary as an innovative tool in soundscape evaluation. Brigitte Schulte-Fortkamp (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany, brigitte.schulte-fortkamp@tu-berlin.de) and Klaus Genuit (HEAD acoustics GmbH, D-52120 Herzogenrath, Germany)

A new field study evaluating soundscapes investigates closely the reactions of traffic noise with a particular regard to the street surface. The combination of methods with different sensibilities for the subject's process of perceiving and evaluating noise in such ambiences is necessary for a reliable and valid analysis and interpretation of data. Acoustic measurements are carried out in critical segments of the street as well as in the respective apartments of the inhabitants, which are questioned in narrative interviews. The acoustic measurements are taken simultaneously in the apartment and on the street. Apartments were selected which issue into the street; outside measurements are performed in front of the buildings on the sidewalk. During the interviews in the apartments the occurring noises are registered by noisebook. As a rule the measurement spot within the apartment is the area in which the interviewed person mostly resides, when he/she takes repose. Further analysis points out the importance of an extended evaluation with an acoustical diary which combines technical and sociological measurement procedures. Performance of the entire data collection process and first results will be discussed.

3:00–3:15 Break

Contributed Papers

3:15

2pNSa6. House attenuation of aircraft and street traffic noise in relation to annoyance: A pilot study. Karl D. Kryter (Dept. of Commun. Disord., San Diego State Univ., San Diego, CA 92182, kdkryter@earthlink.net)

House barrier sound attenuation was measured as differences between the levels of street traffic and aircraft noise found outdoors near the front (street facing) and rear façades, and in rooms of two one-story houses. Also, attenuation of street traffic and aircraft noise between the front and rear façades of similar one-story houses was calculated in accordance with acoustical models. Findings: (1) relative to levels measured near the front façade, street traffic noise near the rear façade, and in rooms inside the houses, was, on average, 9 max dB and 8 SEL dB more attenuated than was aircraft noise. (2) When attenuation of street traffic noise was measured relative to its level near the front façade, and attenuation for aircraft noise was measured relative to its level near the rear façade, street traffic noise was, on average, 7 max dB and 6 SEL dB more attenuated than aircraft noise. (3) Calculated barrier attenuation of street traffic noise was, on average, 1.3 max dB less than measured; calculated barrier attenuation of aircraft noise differed from measured by ~0.4 max dB. Implications of the findings for guidelines for assessing impact of these noises on communities and land-use zoning are discussed.

3:30

2pNSa7. The effect of spatial distribution on the annoyance caused by simultaneous sounds. Joos Vos, Adelbert W. Bronkhorst (Dept. of Percept. TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands), and Thomas Fedtke (Physikalisch-Technische Bundesanstalt, Bundesallee 100, D-38116 Braunschweig, Germany)

A considerable part of the population is exposed to simultaneous and/or successive environmental sounds from different sources. In many cases, these sources are different with respect to their locations also. In a laboratory study, it was investigated whether the annoyance caused by the multiple sounds is affected by the spatial distribution of the sources. There were four independent variables: (1) sound category (stationary or moving), (2) sound type (stationary: lawn-mower, leaf-blower, and chain saw; moving: road traffic, railway, and motorbike), (3) spatial location (left, right, and combinations), and (4) A-weighted sound exposure level (ASEL of single sources equal to 50, 60, or 70 dB). In addition to the individual sounds in isolation, various combinations of two or three different sources within each sound category and sound level were presented for rating. The annoyance was mainly determined by sound level and sound source type. In most cases there were neither significant main effects of spatial distribution nor significant interaction effects between spatial distribution and the other variables. It was concluded that for rating the spatially distrib-

uted sounds investigated, the noise dose can simply be determined by a summation of the levels for the left and right channels. [Work supported by CEU.]

3:45

2pNSa8. Evaluation of urban soundscape by future architects. Mei Zhang and Jian Kang (School of Architecture, Sheffield Univ., Western Bank, Sheffield S10 2TN, UK, m.zhang@sheffield.ac.uk)

Urban soundscape design has drawn great attention along with the ever increasing urban noise level. Previous research shows that people with different social and demographic backgrounds may have different sound preferences in urban environment. For example, gender, age, and cultural background have been proved to affect peoples soundscape evaluation. In this research the soundscape evaluation by a group of architectural students was made, investigating how the future architectural/urban designers value the urban soundscape they experience everyday, and how they would like to design urban soundscape. The soundscape evaluation/walk was conducted with 60 students in a number of urban open public spaces in Sheffield, UK. The evaluation was both for the overall soundscape and for individual sounds. Eighteen indices with seven-point bipolar rating scale were used. Both connotative meanings of urban environment sounds, such as calming–agitating, interesting–boring and like–dislike, and denotative meanings, such as quiet–noisy, sharp–flat and smooth–rough, were considered. In addition, the students were asked to give design suggestions. The semantic differential technique was applied for the analysis. Comparison was also made between this special group and general public. The results are useful for studying the links/interactions between general architectural/urban design and the soundscape design. [Work supported by British Academy.]

4:00

2pNSa9. Evaluation of the performance of the blast analysis and measurement system. George A. Luz (U.S. Army Ctr. for Health Promotion and Preventive Medicine, Aberdeen Proving Ground, MD 21010-5403, George.Luz@amedd.army.mil)

In the years since the introduction of the C-weighted day–night average sound level (DNL) to assess the noise of military explosives, Army practice has evolved to incorporate linear peak sound-pressure level into the evaluation of military training noise. Although the DNL remains as the method of choice for National Environmental Policy Act (NEPA) documentation and for land-use planning, peak level is used by firing range operators for day-to-day complaint management. Several different monitoring system designs are being used at Army installations to provide range operators with real-time feedback on blast noise levels in nearby residential areas. One of these, the Blast Analysis and Measurement (BLAM) system, is a modified version of a sonic boom monitor designed by the U.S. Air Force. Data collected from two BLAM units located near a 120-mm tank gunnery range were evaluated in terms of hit rate and false-alarm rate over a range of 94 to 140 decibels linear peak. Hit- and false-alarm rates are compared with hit- and false-alarm rates reported for other blast noise monitoring system designs.

TUESDAY AFTERNOON, 25 MAY 2004 LIBERTY 1/2, 3:25 TO 5:00 P.M.

Session 2pNSb

Noise: Topics in Active Noise Control

Scott D. Sommerfeldt, Chair
Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602-4673

Chair's Introduction—3:25

Contributed Papers

3:30

2pNSb1. A linear independence method for system identification/ secondary path modeling for active control. Benjamin Kim and David Swanson (Appl. Res. Lab. and Grad. Prog. in Acoust., The Penn. State Univ., University Park, PA 16802, benkim@psu.edu)

A novel method for noninvasive system identification/secondary path modeling has been developed for single- and multi-channel filtered-x LMS-based active noise control (ANC). The problem of on-line secondary path modeling is recognized as one of linear dependence associated with an underdetermined system, a one-equation/two-unknown problem in which the highly correlated primary source and secondary source contributions to the error signal are not readily distinguishable. The method resolves this uniqueness issue by introducing a second equation with similar unknowns. The critical linear independence of the two equations, hence the proposed designation, is achieved with a single perturbation of the control filter output thereby making the system solvable. This new secondary path modeling strategy was implemented using a novel real-time DSP control architecture and tested on a novel transducer-less system devised to investigate the behavior of ANC algorithms. Results of narrow-band, broadband, and multi-channel tests reveal estimates of exceptional accuracy, in both magnitude and phase; any bias in the secondary path estimate, whether introduced by the primary noise source or other secondary sources in the system, is eliminated using this method. Its quickness as well contributes significantly to the stability and performance of filtered-x LMS-based controllers. [Work supported by ONR.]

3:45

2pNSb2. Hybrid feedforward-feedback active noise control for circumaural headsets. Alexander D. Streeter, Laura R. Ray, and Robert D. Collier (Thayer School of Eng., Dartmouth College, Hanover, NH 03755)

Traditional stability-performance tradeoffs pose limitations on active noise reduction (ANR) using feedback control, which are evident in circumaural communication headsets. In these systems, the cavity resonant behavior necessitates low feedback gains, redcing performance. Feedforward ANR using a Lyapunov-tuned least-mean-square filter dramatically enhances noise reduction performance compared with feedback ANR [Cartes *et al.*, J. Acoust. Soc. Am. **111**, 1758–1771 (2002); Collier *et al.*, NOISECON (2003)]. However, feedforward performance is sensitive to the noise source stationarity, and the frequency-dependent forward path gain reduces stability margins. This paper presents experimental results for a hybrid feedforward-feedback ANR system, which enhances performance

and gain margins for both stationary and nonstationary noise. Algorithms are optimized and measurements are performed with Thayer's rapid prototyping system and associated low-frequency acoustic test cell using a circumaural hearing protector. In the frequency range 50–800 Hz, the hybrid system provides an average of 27 dB active noise reduction (40 dB total) for tonal noise and 17 dB reduction (32 dB total) for nonstationary noise. Performance below 100 Hz improves by as much as 15 dB over that of individual control components, and gain margin of the hybrid system improves substantially over individual feedforward or feedback components.

4:00

2pNSb3. Active noise control of small axial cooling fans. Brian B. Monson, Scott D. Sommerfeldt, and Connor Duke (Brigham Young Univ., N283 ESC, Provo, UT 84602, bbm9@email.byu.edu)

An active noise control system has been developed for the reduction of noise radiated by small axial cooling fans, such as those found in desktop computers. This system, based on a modified version of the filtered-x LMS algorithm, uses four small actuators surrounding the fan in a mock computer casing. Global attenuation of the fan noise has been demonstrated using four-channel control. Due to industry volume constraints, a smaller system is required to replace the existing one, while still maintaining similar performance. The system is designed with a smaller fan with small actuators, so as not to exceed the size of the original fan. In order to maintain comparable air flow, the smaller fan must run at higher speeds, causing potentially greater output noise levels at higher frequencies. Global attenuation of these levels is desired. Implementation and experimental results of the system will be presented. A comparison of the two systems, and the feasibility of the replacement of the larger fan by the active control system with a smaller fan will be discussed.

4:15

2pNSb4. Aircraft interior ANC with light-weight actuators. Christian Gerner, Delf Sachau (Helmut-Schmidt-Univ., Mechatronics, Holstenhofweg 85, D-22043 Hamburg, Germany, gerner@unibw-hamburg.de), and Harald Breitbach (Airbus Germany GmbH, D-21229 Hamburg, Germany)

In propeller-driven aircraft the main source for internal noise is tonal disturbances caused by the propeller blades that are passing the fuselage. In a certain four-propeller military transport aircraft the maximum sound level in the cabin can reach up to 110 dB(A), not taking into account any noise control treatments. Inside the semiclosed loadmaster working station (LMWS) the sound level must be reduced down to 86 dB(A). It is proposed to reach this goal with an active noise control system, because passive solutions are too heavy at low frequencies. The optimal positions of the loudspeakers are found by finite-element calculations. These positions have been realized in a full-scale test bed. In the test bed a reduction of the sound-pressure level of more than 30 dB within a specified volume was achieved at a frequency of 100 Hz. In this test bed hi-fi speakers are used as secondary actuators. These speakers are heavy and of unsuitable geometric dimensions for an aircraft. Therefore, other actuators, e.g., flat panel speakers, will be investigated with respect to the application to a mock-up of the LMWS.

4:30

2pNSb5. Limitations on the performance of active noise control systems due to subjective effects. Gerard Mangiante (Laboratoire de Mecanique et d'Acoustique, CNRS, 31 Chemin Joseph-Aiguier, 13402 Marseille Cedex 20, France, mangiante@lma.cnrs-mrs.fr)

Although numerous descriptions of the physical effects of active noise control can be found in the literature, there are few papers available about the subjective effects of this control on listeners. Due to these effects, a significant reduction in the sound pressure level of a primary noise may result in a negligible decrease in the signal perceived by a human observer. This paper reports on the performance of an active system, expressed in terms of reduction in loudness level. For that purpose, the equivalent loudness levels before and after active control were computed for various test signals. For pure tones, the calculation of the loudness level was based on Robinson and Dadson equal-loudness contours; for complex signals, this calculation used the ISO Method A. It was shown that the reduction in loudness level due to active control is highly dependent of the sound pressure level of the primary noise: For frequencies lower than 1000 Hz, the subjective reduction decreases with increasing primary noise level. The subjective reduction was also calculated as a function of the difference in amplitude and phase between the primary and secondary pressure variations. The reduction expressed in phons was then compared to that expressed in dB.

4:45

2pNSb6. Acoustic feedback cancellation for active noise control in pipe. Yun-Hui Liu and Chuan-Yu Hung (Southern Taiwan Univ. of Technol., 1 Nan-Tai St., Yung-Kang City, Tainan Hsien, Taiwan, ROC 710, yhliu@mail.stut.edu.tw)

In most practical applications of active noise control, acoustic feedback is a major problem that often interferes with the operation of the control system and even renders it unstable. The optimal collocated positions of secondary source and two microphones are studied and the theoretical equations that reduce the influence of acoustic feedback are developed based on plane-wave transmission theory of sound in a duct. The original signal of primary noise is obtained from the measured signals of reference microphone and error microphone, which are transformed to frequency domain by FFT and operated. In this study, the performance of the proposed method is compared with other traditional time-domain methods by means of simulation and analysis using LABVIEW programming language. The results showed that the proposed method can effectively reduce the influence of acoustic feedback. In practical application of active noise control in a duct, the adaptive control or other control method incorporating the proposed method in this paper will improve the stability and performance effectively.

Session 2pPP

Psychological and Physiological Acoustics: Auditory Periphery and Models

Laurel H. Carney, Chair
Institute for Sensory Research, Syracuse University, 621 Skytop Road, Syracuse, New York 13244-5290

Contributed Papers

1:30

2pPP1. Coupling of earphones to human ear and to coupler. Dejan Ciric (Faculty of Electron. Eng., Beogradska 14, 18000 Nis, Serbia and Montenegro, dciric@elfak.ni.ac.yu) and Dorte Hammershoi (Aalborg Univ., DK-9220 Aalborg O, Denmark)

The use of a standardized acoustical coupler should enable a calibration of audiometric earphones which ensures that the thresholds determined in the audiometry will be independent of the earphone type. This requires that the coupler approximates the average human ear closely. Nevertheless, the differences among earphones as well as between human ears and the coupler affect the results of audiometric measurements inducing uncertainty. The influence of these differences is examined by investigating the coupling of different earphones to human ears and to the standardized coupler. This is done by measurement of the transfer functions from input voltage of the earphone terminals to the entrance of the ear canal in two situations: (1) open, and (2) blocked. Similar measurements were carried out with the coupler, but since the ''ear-canal entrance'' is not well-defined for the coupler, the mentioned measurements were done at different depths in the coupler. The earphone's coupling to (i) human ears and to (ii) the coupler, described in terms of the pressure division at the entrance of the ear canal, were compared. The results indicate that the coupling to the human ear and the coupling to the standardized coupler differ.

1:45

2pPP2. Using reflectance phase to estimate the acoustic response at the tympanic membrane. Douglas H. Keefe (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, keefe@boystown.org)

A common problem in ear-canal acoustics is to estimate the SPL and acoustic transfer functions (ATFs) at the tympanic membrane (TM) based on a response measured at a distal probe location, but standing waves complicate interpretation above 1.5 kHz. Transmission is well approximated up to 8 kHz by modeling the ear canal as a rigid-walled cylindrical tube; the model transforms measurements at the probe tip to those at the TM. Based on SPL and ATFs measured at the probe in 42 normal-hearing and 18 conductive-impaired ears from 0.25–8 kHz, the equivalent SPL and ATFs at the TM are calculated for ambient and pressurized conditions in the ear canal. Tube area is calculated based on probe-tip size, while tube length from probe to TM is calculated from the group delay of the reflectance phase averaged from 6 to 8 kHz. In normal-hearing ears, the mean ($\pm$standard deviation) of the SPL difference at the TM *re*: probe has a maximum of 9.2 ± 3.8 dB at 3.7 ± 0.8 kHz. The equivalent TM admittance has a wideband shape similar to *in vivo* measurements of umbo velocity level at constant SPL at the TM [Ruggero and Temchin, 2003]. [Work supported by NIH.]

2:00

2pPP3. The impact of cochlear fine structure on hearing thresholds and DPOAE levels. Jungmee Lee, Glenis Long (Speech and Hearing Sci., Grad. Ctr. of CUNY, 365 5th Ave., New York, NY 10016, jmlee6@msn.com), and Carrick L. Talmadge (The Univ. of Mississippi, University, MS 38677)

Although otoacoustic emissions (OAE) are used as clinical and research tools, the correlation between OAE behavioral estimates of hearing status is not large. In normal-hearing individuals, the level of OAEs can vary as much as 30 dB when the frequency is changed less than 5%. These pseudoperiodic variations of OAE level with frequency are known as fine structure. Hearing thresholds measured with high-frequency resolution reveals a similar (up to 15 dB) fine structure. We examine the impact of OAE and threshold fine structures on the prediction of auditory thresholds from OAE levels. Distortion product otoacoustic emissions (DPOAEs) were measured with sweeping primary tones. Psychoacoustic detection thresholds were measured using pure tones, sweep tones, FM tones, and narrow-band noise. Sweep DPOAE and narrow-band threshold estimates provide estimates that are less influenced by cochlear fine structure and should lead to a higher correlation between OAE levels and psychoacoustic thresholds. [Research supported by PSC CUNY, NIDCD, National Institute on Disability and Rehabilitation Research in U.S. Department of Education, and The Ministry of Education in Korea.]

2:15

2pPP4. Can tone-burst otoacoustic emissions be used to measure basilar-membrane response in humans? Michael Epstein, Søren Buus (ECE Dept. (440 DA), Inst. of Hearing, Speech and Lang. and Commun. and DSP Ctr., Northeastern Univ., Boston, MA 02115, mepstein@ece.neu.edu), and Mary Florentine (Northeastern Univ., Boston, MA 02115)

Tone-burst otoacoustic emissions (TBOAEs) are a potential tool for examining basilar-membrane activity in humans. However, their use requires knowledge of just how the relation between TBOAEs and basilar-membrane input-output (BM I/O) activity depends on measurement and analysis parameters. The present experiment examined three such parameters: (1) response-window timing, (2) response delinearization, and (3) stimulus frequency. Varying the onset of the post-stimulus-recording analysis window had little effect on the function relating emission amplitude to SPL. Response delinearization, if anything, tended to obscure the relation between TBOAE and stimulus level. Changes in stimulus frequency over a range of 1 Bark had no effect on TBOAE amplitude. Although none of these experimental manipulations had a significant effect on TBOAE measurements, a judicial choice of the experimental parameters can optimize the assessment of BM I/O functions. It is concluded that TBOAEs provide a useful tool for objective BM I/O function analysis in humans. [Work supported by NIH-NIDCD Grant R01DC00187.]

2:30

2pPP5. Relating psychophysical and otoacoustic emission suppression tuning curves. Jeffrey J. DiGiovanni and Erika M. Zettner (School of Hearing Speech and Lang. Sci., Ohio Univ., Athens, OH 45701, digiovan@ohio.edu)

Psychophysical tuning curves (PTCs) are a well established method for estimating frequency selectivity in humans. PTCs are closely related to the shape of neural tuning curves. Furthermore, these two measurements are believed to represent frequency resolution processes primarily occurring in the cochlea. More recently, otoacoustic emission suppression tuning curves (STCs) have been measured and found to be characteristically similar in shape and tip frequency to PTCs (Abdala *et al.*, 1996). STCs may represent an objective, noninvasive, measure of initial stages of cochlear tuning. The goal of this study was to compare and quantify the difference between transient evoked STCs (TE-STC) and PTCs. TE-STCs and PTCs were obtained in each normal-hearing subject with 4000-Hz probe tones. Preliminary data show that Q10 values and tip levels follow the same trend for both tuning curves. However, actual values of these parameters differ. Implications of methodological differences between these two paradigms will be discussed.

2:45

2pPP6. On the possible role of MOC efferents in speech reception in noise. Oded Ghitza (Sensimetrics Corp., 48 Grove St., Somerville, MA 02144, oded@sens.com)

The hypothesis explored in this study is that the MOC efferent system plays an important role in speech reception in the presence of sustained background noise. This talk describes efforts to assess this hypothesis using a test of initial consonant reception (the Diagnostic Rhyme Test) performed by subjects with normal hearing. Activation of selected parts of the efferent system was attempted by presenting speech and noise in various configurations (gated/continuous, monaural/binaural). Initial results of these experiments show a gated/continuous difference analogous to the ''masking overshoot'' in tone detection. These results are interpreted to support the hypothesis of a significant efferent contribution to initial phone discrimination in noise. [Work supported by AFOSR.]

3:00

2pPP7. Revisiting the Fano factor in peripheral neurons. Willy Wong (Dept. of Elec. and Computer Eng., and Inst. of Biomaterials and Biomed. Eng., Univ. of Toronto, ON M5S 3G4, Canada, willy@eecg.utoronto.ca)

Despite the wealth of studies on the dynamic characteristics of peripheral auditory neurons, very little has been reported on the higher statistical moments of the neural spike train. The notable exception is the study by Teich and Khanna (1985) where both the mean and the variance of the neural count are reported. The simplest model one can ascribe to a neural spike train is a homogeneous Poisson process. However, experimental data do not bear out such predictions. Other models have been proposed but the general consensus is that the underlying process is far from simple. We offer an alternative account of the fluctuations that occur at the peripheral level. Our explanation does not rely on assumptions regarding the process underlying individual spikes. Instead, we make use of the information-theoretical model of the neuron that we have been developing over the past 10 years (Norwich and Wong, 1995; Wong, 1997). The two key results predicted by the model are that (a) the mean-variance ratio has an approximate value of 2 and (b) the distribution governing the neural count is Gaussian to a good approximation.

3:15

2pPP8. Towards an analytical derivation of a computational auditory masking model. Joerg M. Buchholz (Univ. of Western Sydney, MARCS Auditory Labs., Locked Bag 1797, Penrith South DC NSW 1797, Australia, j.buchholz@uws.edu.au)

Due to the fact that the signal processing underlying auditory masking is highly nonlinear, existing auditory masking models are usually designed on rather a trial-and-error basis than in a well-defined analytical way. Employing a model structure, which is based on the concept of signal-dependent compression (SDC), a method is presented which demonstrates the analytical derivation of a computational auditory masking model. Given the SDC-based model approach, analytical approximations of the model's masked threshold simulations can be derived, which can directly be compared to masking functions known from the relevant literature. In this way, the present model structure has been adjusted in order to describe mainly two masking functions: (i) Webber's law, which approximates the masker level dependency of simultaneous noise-on-tone masking, and (ii) a function proposed by Jestead *et al.*, J. Acoust. Soc. Am. **71**, 950–962 (1982), which approximates the masker level and test-signal delay dependency of forward masking. The practical applicability of the derived model structure has been verified by comparing model simulations to various known psychoacoustical data of simultaneous masking and forward masking for broadband noise masking tones.

3:30

2pPP9. Modeling the detection and level discrimination of tones in the presence of masking noise. Xuedong Zhang (Dept. of Biomed. Eng. and Hearing Res. Ctr., Boston Univ., Boston, MA 02215, zxd@bu.edu) and Laurel H. Carney (Syracuse Univ., Syracuse, NY)

The performance of three model detectors was evaluated for masked detection and level discrimination of tones in noise with different bandwidths and in a roving level paradigm (Kidd *et al.*, 1989). The model detectors were based on (1) the output energy of several auditory filters tuned to different frequencies, (2) the envelope statistics (peakiness) of the auditory filter outputs, and (3) the cross-correlations between auditory-nerve (AN) fibers in a population model. The energy-based model predicted detection thresholds, but failed to predict masked level discrimination; this detector was robust for roving-levels in the wideband condition by combining information from different auditory filter outputs and the prediction was affected by roving narrow-band masker levels. Thresholds of the envelope-based detector were worse than human thresholds for some conditions; however, this model was robust in roving-levels for both wideband and narrow-band maskers. The monaural cross-correlation detector included model cells that were sensitive to temporal cues in model AN inputs in response to wideband noise and also included cells that were sensitive to input level changes. This model naturally combined rate and temporal information and predicted performance for both masked detection and discrimination. [Work supported by NIDCD R01-01641.]

3:45

2pPP10. Physiologic correlates to background noise acceptance. Joanna Tampas, Ashley Harkrider, and Anna Nabelek (Dept. of Audiol. and Speech Pathol., Univ. of Tennessee, 457 S. Stadium Hall, Knoxville, TN 37996)

Acceptance of background noise can be evaluated by having listeners indicate the highest background noise level (BNL) they are willing to accept while following the words of a story presented at their most comfortable listening level (MCL). The difference between the selected MCL and BNL is termed the acceptable noise level (ANL). One of the consistent findings in previous studies of ANL is large intersubject variability in acceptance of background noise. This variability is not related to age, gender, hearing sensitivity, personality, type of background noise, or speech perception in noise performance. The purpose of the current experiment was to determine if individual differences in physiological activ-

ity measured from the peripheral and central auditory systems of young female adults with normal hearing can account for the variability observed in ANL. Correlations between ANL and various physiological responses, including spontaneous, click-evoked, and distortion-product otoacoustic emissions, auditory brainstem and middle latency evoked potentials, and electroencephalography will be presented. Results may increase understanding of the regions of the auditory system that contribute to individual noise acceptance.

TUESDAY AFTERNOON, 25 MAY 2004 LIBERTY 1/2, 1:30 TO 3:10 P.M.

Session 2pSA

Structural Acoustics and Vibration and Noise: Urban Transit Noise

Paul L. Burge, Cochair
Acentech, 33 Moulton Street, Cambridge, Massachusetts 02138

Daniel R. Raichel, Cochair
2727 Moore Lane, Fort Collins, Colorado 80526

Chair's Introduction—1:30

Invited Papers

1:35

2pSA1. New developments in transit noise and vibration criteria. Carl E. Hanson (Harris Miller Miller & Hanson, Inc., 15 New England Executive Park, Burlington, MA 01803, chanson@hmmh.com)

Federal Transit Administration (FTA) noise and vibration impact criteria were developed in the early 1990's. Noise criteria are ambient-based, developed from the Schultz curve and fundamental research performed by the U.S. Environmental Protection Agency in the 1970's. Vibration criteria are single-value rms vibration velocity levels. After 10 years of experience applying the criteria in assessments of new transit projects throughout the United States, FTA is updating its methods. Approach to assessment of new projects in existing high-noise environments will be clarified. Method for assessing noise impacts due to horn blowing at grade crossings will be provided. Vibration criteria will be expanded to include spectral information. This paper summarizes the background of the current criteria, discusses examples where existing methods are lacking, and describes the planned remedies to improve criteria and methods.

1:55

2pSA2. Vibration criteria for transit systems in close proximity to university research activities. Steven Wolf (Parsons Brinckerhoff, 505 S. Main St., Orange, CA 92868)

As some of the newer LRT projects get closer to research facilities the question arisesi "how do you assess the potential impact of train operations on the activities within these types of facilities?" There are several new LRT projects that have proposed alignments near or under university research facilities. The traditional ground vibration analysis at these locations is no longer valid but requires a more sophisticated approach to identifying both criteria and impact. APTA, ISO, IES, and FTA vibration criteria may not be adequate for the most sensitive activities involving single cell and nano technology research. The use of existing ambient vibration levels is evaluated as a potential criteria. A statistical approach is used to better understand how the train vibration would affect the ambient vibration levels.

2:15

2pSA3. Determination of noise and vibration impacts of construction and operation of the Second Avenue Subway. Jessica L. Pristera (AKRF, Inc., 117 E. 29th St., New York, NY 10016, jessica_pristera@akrf.com)

An acoustical study was conducted to determine the potential for airborne noise and ground-borne noise and vibration impacts generated by construction and operation of the Second Avenue Subway. The study was performed in support of an environmental impact statement (EIS) that defined the areas along the proposed Second Avenue Subway corridor where any significiant impacts would occur as a result of construction activity and operation of the Second Avenue Subway. Using FTA guideline procedures, project-generated noise levels from subway construction and operations were determined. Construction noise levels exceeded operational noise levels. With limited alternative construction methods, practical mitigation methods were determined to reduce impacts.

2:35

2pSA4. Update on Chicago Transit Authority noise, how the city is attacking the problem. Matthew McDuffee and Dominique Cheenne (Columbia College Chicago, 600 S. Michigan Ave., Chicago, IL 60605)

Chicago's light rail rapid transit system is known as the L. The city of Chicago has the oldest standing tracks in the country with some portions of the system dating back to the 1880s. The noise of the L is a signature mark of Chicago and has long been dealt with by residents due to the integral nature of the transportation system to the city. The paper outlines what the CTA (Chicago Transit Authority) planning and development office defines as the four major sources of noise, and how they deal with the issues. Among these are thump, roar, rumble, and squeal. These issues are dealt with through rail grinding programs, wheel truing, and structural improvements, as well as wheel dampening rings, rail lubrication, and reconstruction projects to reduce the steepness of curves. A before and after study of one of the recent CTA construction projects was done to see if these goals are being met by the reconstruction projects being performed in the field. The CTA has agreed to cooperate with me in this project; the Vice President of planning and development has offered data and the CTA approval of the independent study.

Contributed Paper

2:55

2pSA5. Noise annoyance caused by magnetic levitation train passbys. Joos Vos (Dept. of Percept. TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands)

In a laboratory study, the annoyance caused by the passby sounds from a magnetic levitation (maglev) train was investigated. The outdoor A-weighted sound exposure level (ASEL) of the maglev sounds varied from 65 to 90 dB. The driving speed of the maglev train varied from 100 to 400 km/h. Four important results were obtained. Provided that the outdoor ASELs were the same, (1) the annoyance was independent of the driving speed of the maglev train, (2) the annoyance caused by the maglev train was considerably higher than that caused by intercity trains, (3) the annoyance caused by the maglev train was hardly different from that caused by road traffic (passenger cars and trucks), and (4) the results (1)–(3) held true both for open or closed windows. On the basis of the present results, it might be expected that the sounds are equally annoying if the ASELs of the maglev-train passbys are at least 5 dB lower than those of the intercity train passbys. Consequently, the results of the present experiment do not support application of a railway bonus to the maglev-train sounds. Issues for future research, such as exploring further contributions of nonacoustic factors, will be discussed.

TUESDAY AFTERNOON, 25 MAY 2004 NEW YORK BALLROOM B, 1:00 TO 5:15 P.M.

Session 2pSC

Speech Communication: Speech Perception and Production in Monolingual and Bilingual Acquisition

Linda Polka, Cochair
School of Communication Sciences and Disorders, McGill University, Beatty Hall, 1266 Pine Avenue West, Montreal, Quebec H3G 1A8, Canada

Megha Sundara, Cochair
School of Communication Sciences and Disorders, McGill University, Beatty Hall, 1266 Pine Avenue West, Montreal, Quebec H3G 1A8, Canada

Chair's Introduction—1:00

Invited Papers

1:05

2pSC1. Cross-linguistic experiments in word-form recognition. Marilyn Vihman (Univ. of Wales Bangor, Gwynedd LL57 2AS, Wales, UK, m.vihman@bangor.ac.uk)

When do children first represent word forms without experimental training or contextual support? Both English- and Welsh-learning children were tested, replicating Halle and Boysson-Bardies (1994: French, 11 months.). Twelve children acquiring English showed word-form recognition by 11 months (Vihman *et al.*, in press); 12 Welsh children showed the effect at 12 months but a separate sample of 12 tested at 11 months did not (Vihman and DePaolis, 1999). A subsequent study of 16 children using event-related potentials (ERPs) showed word-form recognition within 250 ms for English at 11 months (Thierry *et al.*, 2003). Attempts to locate the age of onset longitudinally proved problematic: Repeated tests of single samples of English and Welsh monolingual children (12 each) at 9, 10, 11, and 12 months showed that infant episodic memory interferes sufficiently with longitudinal observation based on a single set of stimuli to preclude drawing any conclusions. Cross-sectional samples of monolingual English and Welsh children (24 each) are currently being tested at 9 to12 months, using both head turn and ERPs, as are English/Welsh bilingual children at 11 months. These studies should yield solid information as to the age of onset of spontaneous word form representation. [ESRC support is gratefully acknowledged.]

1:25

2pSC2. Native sound category formation in simultaneous bilingual acquisition. Laura Bosch (Dept. of Psicologia Basica, Univ. of Barcelona, P. Vall d'Hebron, 171, 08035 Barcelona, Spain, laurabosch@ub.edu)

The consequences of early bilingual exposure on the perceptual reorganization processes that occur by the end of the first year of life were analyzed in a series of experiments on the capacity to discriminate vowel and consonant contrasts, comparing monolingual and bilingual infants (Catalan/Spanish) at different age levels. For bilingual infants, the discrimination of target vowel contrasts, which reflect different amount of overlapping and acoustic distance between the two languages of exposure, suggested a U-shaped developmental pattern. A similar trend was observed in the bilingual infants discrimination of a fricative voicing contrast, present in only one of the languages in their environment. The temporary decline in sensitivity found at 8 months for vowel targets and at 12 months for the voicing contrast reveals the specific perceptual processes that bilingual infants develop in order to deal with their complex linguistic input. Data from adult bilingual subjects on a lexical decision task involving these contrasts add to this developmental picture and suggest the existence of a dominant language even in simultaneous bilingual acquisition. [Work supported by JSMF 10001079BMB.]

1:45

2pSC3. Speech perception and production in bilinguals and monolinguals: Convergent and divergent patterns in adulthood and development. Megha Sundara and Linda Polka (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada)

Most investigations of the role of language experience have looked at monolingual subjects. Data are presented from a series of experiments with consonant contrasts to highlight how bilingual adults exposed to Canadian-English and Canadian-French from birth accommodate two sets of input regularities in their perception and production behavior. Results from these investigations indicate that there are similarities as well as differences between adult simultaneous bilinguals and their monolingual peers. Recent findings from our lab exploring phonetic discrimination and word segmentation in monolingual and bilingual infants and children, which provide insights into the developmental roots of these adult patterns, are also presented. The findings from this work illustrate that during the developmental process bilingual infants and children differ from their monolingual peers in rate as well as pattern of development for some perceptual tasks but not others.

2:05

2pSC4. Developmental and communicative factors affecting VOT production in English and Arabic bilingual and monolingual speakers. Ghada Khattab (Speech and Lang. Sci., King George VI Bldg., Newcastle Univ., Queen Victoria Rd., Newcastle upon Tyne NE1 7RU, UK, ghada.khattab@ncl.ac.uk)

VOT patterns were investigated in the production of three Lebanese–English bilinguals' aged 5, 7, and 10, six aged-matched monolingual controls from the bilinguals' immediate communities, and the parents of bilinguals and monolinguals. The aim was to examine the extent to which children exposed to two languages acquire separate VOT patterns for each language and to determine the factors that affect such acquisition. Results showed that VOT patterns for each bilingual child differed significantly across the two languages. But while the contrast in English resembled a monolingual-like model, that for Arabic exhibited persisting developmental features; explanations were offered in terms of the relationship between input and complexity of voicing lead production. Evidence was used from developmental changes that were noted for two of the bilingual subjects over a period of 18 months. English code-switches produced by the bilinguals during Arabic sessions exhibited different VOT patterns from those produced during English sessions, which underlined the importance of taking the language context into consideration. Finally, results from monolinguals and bilinguals showed that the short lag categories for the two languages were different despite a degree of overlap. Such findings require finer divisions of the three universal VOT categories to account for language-specific patterns.

2:25

2pSC5. Acquisition of English word stress patterns in early and late bilinguals. Susan G. Guion (Linguist. Dept., 1290 Univ. of Oregon, Eugene, OR 97403, guion@uoregon.edu)

Given early acquisition of prosodic knowledge as demonstrated by infants' sensitivity to native language accentual patterns, the question of whether learners can acquire new prosodic patterns across the life span arises. Acquisition of English stress by early and late Spanish–English and Korean–English bilinguals was investigated. In a production task, two-syllable nonwords were produced in noun and verb sentence frames. In a perception task, preference for first or last syllable stress on the nonwords was indicated. Also, real words that were phonologically similar to the nonwords were collected. Logistic regression analyses and ANOVAs were conducted to determine the effect of three factors (syllable structure, lexical class, and stress patterns of phonologically similar words) on the production and perception responses. In all three groups, stress patterns of phonologically similar real words predicted stress on nonwords. For the two other factors, early bilinguals patterned similarly to the native-English participants. Late Spanish–English bilinguals demonstrated less learning of stress patterns based on syllabic structure, and late Korean–English bilinguals demonstrated less learning of stress patterns based on lexical class than native-English speakers. Thus, compared to native speakers, late bilinguals' ability to abstract stress patterns is reduced and affected by the first language. [Work supported by NIH.]

2:45–3:00 Break

3:00

2pSC6. Neural network simulation of habituation and dishabituation in infant speech perception. Bruno Gauthier, Rushen Shi, and Robert Proulx (Dept. of Psych., Univ. of Quebec at Montreal, Montreal, QC H3C 3P8, Canada, t16025@internet.uqam.ca)

The habituation techniques used in infant speech perception studies are based on the fact that infants show renewed interest towards novel stimuli. Recent work has shown the possibility of using artificial neural networks to model habituation and dishabituation (e.g., Schafer and Mareschal, 2001). In our study we examine weather the self-organizing-feature-maps (SOM) (Kohonen, 1989) are appropriate for modeling short-term habituation to a repeated speech stimulus. We found that although SOMs are particularly useful for simulating categorization, they can be modified to model habituation and dishabituation, so that they can be applied to direct comparisons with behavioral data on infants' speech discrimination abilities. In particular, we modified the SOMs to include additional parameters that control the relation of input similarity, lateral inhibition, and local and lateral activation between neurons. Preliminary results suggest that these parameters are sufficient for the network to simulate the loss of sensitivity of the auditory system due to the presentation of multiple tokens of a speech stimulus, as well as to model the recovery of sensitivity to a novel stimulus. The implications of this approach to infant speech perception research will be considered.

3:15

2pSC7. Phonetic representation of frequent function words in 8-month-old infants. Rushen Shi (Dept. of Psych., Univ. of Quebec at Montreal, Montreal, QC H3C 3P8, Canada, shi.rushen@uqam.ca), Janet Werker (Univ. of BC, Vancouver, BC V6T 1Z4, Canada), and Anne Cutler (Max Planck Inst. for Psycholinguist., Nijmegen, The Netherlands)

Recent work by a number of researchers showed that even preverbal infants detect and recognize functors in continuous speech. In our research, English-learning infants aged 11 to 13 months, but not 8 months, recognized frequent and infrequent functors as a class, and represented them in segmental detail (Shi *et al.*, 2003; Shi *et al.*, 2004). Here we report a study on 8-month-old infants' recognition and representation of high versus low frequency functors. Infants heard sequences containing a lexical word preceded by a high frequency functor ''the,'' versus a nonsense functor ''kuh,'' differing from ''the'' only in the initial consonant, with the prosody unchanged. Another group of 8-month-olds heard sequences containing a lexical word preceded by a low frequency functor ''its,'' versus a nonsense functor ''ots.'' Recognition of functors would be indicated by longer listening time to sequences containing real functors. Results reveal no differential listening time between ''the+lexical word(LW)'' and ''kuh+LW,'' nor between ''its+LW'' and ''ots+LW;'' however, ''the+LW'' and ''kuh+LW'' together induced longer listening time than ''its+LW'' and ''ots+LW.'' We conclude that 8-month olds recognize the frequent, familiar ''the'' in continuous speech, but it is underspecified phonetically in infants' initial lexicon. Our previous work indicates detailed specification by 11 months.

3:30

2pSC8. Segments and segmental properties in cross-language perception: Korean perception of English obstruents in various prosodic locations. Kenneth de Jong, Noah Silbert, and Hanyong Park (Dept. of Linguist., 322 Memorial Hall, Indiana Univ., Bloomington, IN 47405, kdejong@indiana.edu)

Experimental models of cross-language perception and second-language acquisition (such as PAM and SLM) typically treat language differences in terms of whether the two languages share phonological segmental categories. Linguistic models, by contrast, generally examine properties which cross classify segments, such as features, rules, or prosodic constraints. Such models predict that perceptual patterns found for one segment will generalize to other segments of the same class. This paper presents perceptual identifications of Korean listeners to a set of voiced and voiceless English stops and fricatives in various prosodic locations to determine the extent to which such generality occurs. Results show some class-general effects; for example, voicing identification patterns generalize from stops, which occur in Korean, to nonsibilant fricatives, which are new to Korean listeners. However, when identification is poor, there are clear differences between segments within the same class. For example, in identifying stops and fricatives, both point of articulation and prosodic position bias perceptions; coronals are more often labeled fricatives, and syllable initial obstruents are more often labeled stops. These results suggest that class-general perceptual patterns are not a simple consequence of the structure of the perceptual system, but need to be acquired by factoring out within-class differences.

3:45

2pSC9. Cross-language perceptual category mapping: Korean perception of English obstruents. Hanyong Park, Kenneth de Jong, and Noah Silbert (Dept. of Linguist., 322 Memorial Hall, Indiana Univ., Bloomington, IN 47405, hanypark@indiana.edu)

Models, such as SLM and PAM, predict that performance on second language sounds is determined by the perceptual relationship of the sounds to the original language categories. To measure this relationship, Schmidt [J. Acoust. Soc. Am. **99**, 3201–3211] had native Korean speakers classify English consonant productions into Korean orthographic categories, and assess the similarity of the consonants to the chosen categories. The current experiment further examines how Korean labeling relates to accuracy in using English orthographic categories. Results show Koreans poorly identify sounds rated as dissimilar from Korean categories. Similarly, sounds that are inconsistently labeled with Korean labels are less accurately identified. These results suggest that accuracy varies with the nearness of the English and a Korean category, and thus English categories are developed from original Korean categories. However, other results indicate that sounds straddling Korean categories can be very accurately labeled, suggesting a complete suppression of native contrasts. In addition, identification accuracy for subjectively odd sounds can be very high, suggesting the development of a new category. Similarly, cases in which both Korean and English labeling is inconsistent show no relationship between the Korean and the English labels, indicating that the English categories are constructed apart from the existing Korean categories.

4:00

2pSC10. Production of Mandarin tones by 3-year-old children acquiring L1 (Mandarin) in an L2 (English) environment. Puisan Wong and Richard G. Schwartz (Ph.D. Prog. in Speech and Hearing Sci., The Grad. School and Univ. Ctr., City Univ. of New York, 365 Fifth Ave., New York, NY 10016)

Thirteen monolingual Mandarin-speaking children residing in the U.S. were recruited to examine their production of the four Mandarin lexical tones in monosyllabic words. A picture-naming task was used to elicit the children's productions of lexical tones in isolated words and in sentence final position. Four mothers were asked to say the same set of words to their children in a picture reading activity. The children's and the mothers' productions were recorded and low-pass filtered at 500 and 400 Hz, respectively, to eliminate phonemic and semantic information. Ten Mandarin-speaking judges were recruited to identify the children's and adults' tone productions from the filtered stimuli. Contrary to the findings of L1 research conducted in countries where Mandarin is the language of the environment, the present results revealed that the lexical tones produced by 3-year-old children acquiring Mandarin as their first language in the U.S. were not yet adultlike. Children's tone productions were more difficult to categorize than the mothers' productions. The judges had significantly more difficulty identifying children's dipping tones than the children's level tones, rising tones, or falling tones, suggesting that the dipping tone posed the most difficulties for the children.

4:15

2pSC11. Language preference in monolingual and bilingual infants. Ayasha Valji and Linda Polka (School of Commun. Sci. and Disord., McGill Univ., 1266 Pine Ave., Montreal, QC H3G 1A8, Canada linda.polka@mcgill.ca)

Previous research shows that infants being raised in single-language families have some basic language discrimination abilities at birth, that these skills improve over the first 6 months of life, and that infants are attending to the rhythmic properties of language to perform these skills. Research has also revealed that newborns and older babies from monolingual families prefer listening to their native language over an unfamiliar language. Data on language discrimination and preference in bilingual infants is very limited but is necessary to determine if the patterns and rate of bilingual language development parallel those of monolingual development, or if exposure to more than one language modifies developmental patterns. The present study addresses this issue by comparing language preference in monolingual English, monolingual French, and bilingual English–French infants between 3 and 10 months of age. Infant preference to listen to passages in three rhythmically different languages (English, French, Japanese) was assessed using a visual fixation procedure. Passages were produced by three female native speakers of each language. Findings will show how native language preference is affected by age and language experience in infants who experience monolingual and bilingual language exposure.

4:30

2pSC12. Syllable-final fricatives: Dutch and English listeners' processing of voicing. Mirjam Broersma (Max Planck Inst. for Psycholinguist., Postbus 310, 6500 AH Nijmegen, The Netherlands, mirjam.broersma@mpi.nl)

English and Dutch have both voiced and voiceless fricatives, and, in English, both occur in either syllable-initial or -final position. In Dutch, however, only voiceless fricatives can occur in syllable-final position. A categorization experiment investigated the processing of the voicing distinction in English fricatives by Dutch and English listeners, and in particular whether Dutch listeners can distinguish voicing contrasts in syllable-final position, and whether preceding vowel length informs their voicing judgments. Dutch and English participants heard tokens from an 11-step /v-f/ continuum, at the end of a single nonword. As the nonword context was kept constant within each block, vowel length could not serve as a cue. Half of the participants heard nonword contexts which were originally pronounced with a final /v/, the other half with final /f/. Therefore, in both cases a mismatch occurred for a subset of the items between vowel length and the other information in the signal. This mismatch hampered the performance of the English but not the Dutch listeners, so that the Dutch listeners outperformed the English in mismatching cases. This suggests that the Dutch did not rely on vowel length as a cue to voicing as strongly as the English.

4:45–5:15
Panel Discussion

TUESDAY AFTERNOON, 25 MAY 2004 RIVERSIDE SUITE, 1:25 TO 4:15 P.M.

Session 2pSP

Signal Processing in Acoustics and Underwater Acoustics: Acoustic Communications—Algorithms and Performance Analysis

Milica Stojanovic, Chair
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Chair's Introduction—1:25

Contributed Papers

1:30

2pSP1. Performance predictions for acoustic communication using passive phase conjugation. Daniel Rouseff (Appl. Phys. Lab., College of Ocean and Fishery Sci., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105)

Passive phase conjugation processing for acoustic communication combines temporal matched filtering with spatial diversity. An efficient decision-directed procedure has been implemented to calculate the matched filters and has been applied successfully to experimental data. In the present paper, a model is developed for the expected mean-squared error in the demodulation output. A normal mode expansion is used for the phase-encoded acoustic field. The matched filter integration is evaluated using a stationary phase approximation. This intermediate result is interpreted using the waveguide invariant. After combining across the array and sampling, formal statistical averages of the demodulation output are calculated. Results are parametrized in terms of the number of array elements and their spatial distribution. The effects of truncating the matched filters and reducing the bandwidth are also quantified. Performance predictions are shown to be in good agreement with experimental results. [Work supported by ONR under the ARL program.]

1:45

2pSP2. Performance comparisons between passive-phase conjugation and decision-feedback equalizer for underwater acoustic communications. T. C. Yang (Naval Res. Lab. Code 7120, 4555 Overlook Ave., Washington, DC 20375)

Passive-phase conjugation (PPC) uses passive time reversal to remove intersymbol interferences (ISI) for acoustic communications in a multipath environment. It is based on the theory of signal propagation in a waveguide, which says that the Green function (or the impulse response function) convolved with its time-reversed conjugate, summed over a large aperture vertical array of receivers is a delta function in space and time. A decision feedback equalizer (DFE) uses a nonlinear filter to remove ISI based on the minimum mean square errors (MSE) between the estimated and the true (or decision) symbols. These two approaches are motivated by different principles. In this paper, we analyze both using a common framework, illustrating the commonality and differences, pro and con between the two methods, and compare their performances in realistic ocean environments. The performance measures are MSE, output signal-to-noise ratio and bit error rate (BER) as a function of the number of receivers. For a small number of receivers, DFE outperforms PPC in all measures. As the

number of receivers increases the BER for both processors approaches zero, but at a different rate. The results are supported by both simulated and real data. [Work supported by ONR.]

2:00

2pSP3. Robust adaptive decision feedback equalization using residual prediction errors. James C. Preisig (Dept. of Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Channel estimated based adaptive decision feedback equalization (DFE) relies on estimates of the channel impulse response and interfering noise field to calculate optimal equalizer filter coefficients. While the DFE offers superior performance to either passive time-reversal or minimum mean squared error (mmse) linear equalization when good quality channel estimates are available, the DFE is more sensitive than the other two techniques to channel estimation errors. A technique incorporating the residual prediction error of the channel estimation algorithm is presented. This algorithm, referred to as the residual prediction error DFE (RPE-DFE), maintains the superior performance that is characteristic of the DFE operating with good quality channel estimates while improving the robustness of the DFE to channel estimation errors. Experimental data is presented that allows for a direct comparison of the RPE-DFE performance with the performance of the traditional DFE, the passive time reversal and mmse linear equalizers, and robust DFEs proposed earlier [M. Stojanovic, J. Proakis, and J. Catipovic, IEEE T-Comm. **43**, 877–886 (1995)]. [Work supported by ONR Ocean Acoustics.]

2:15

2pSP4. Information and data real time transmission acoustic underwater system: TRIDENT. Joel Trubuil, Joel Labat (ENST Bretagne Technopole Brest Iroise CS 83818, 29238 Brest, Cedex 3, France, joel.trubuil@enst-bretagne.fr), and Gerard Lapierre (GESMA, BP42 29240, Brest Naval, France)

The objective of the Groupe d'Etudes Sous-Marines de l'Atlantique (GESMA) is to develop a robust high data rate acoustic link. A real-time receiver recently developed at ENST Bretagne has just been designed to cope with all perturbations induced by such harsh channels. In order to cope with channel features, a spatio-temporal equalizer introduced by J. Labat *et al.* [Brevet FT no. 9914844, ''Perfectionnements aux dispositifs d'galisation adaptative pour recepteurs de systemes de communications numriques,'' Nov. 1999] was recently implemented and evaluated. This equalizer is the core of the receiver platform [Trubuil *et al.*, ''Real-time high data rate acoustic link based on spatio temporal blind equalization: the TRIDENT acoustic system,'' OCEANS 2002]. This paper provides an overview of this project. The context of the study and the design of high data rate acoustic link are presented. Last Brest harbor experiments (2002, 2003) are described. The real time horizontal acoustic link performances are evaluated. Two carriers frequencies are available (20, 35 kHz). Acoustic communications for bit rate ranging from 10 to 20 kbps and for channel length (shallow water) ranging from 500 to 4000 m have been conducted successfully over several hours.

2:30

2pSP5. Reliable Doppler-spread estimation using an evolutionary algorithm for underwater acoustic communications over channels with frequency-selective fading. Jasdeep S. Dhanoa, Richard F. Ormondroyd, and Evan J. Hughes (Dept. of Aerosp., Power and Sensors, Cranfield Univ., The Royal Military College of Sci., Shrivenham, Oxfordshire SN6 8LA, UK)

There is a pressing need for reliable wideband underwater acoustic communication for enhanced performance of diver communications, underwater networks, control links to ROVs and data links to remote sensors and AUVs. The performance of such links is compromised by the doubly-spread acoustic channel, where multipath time-delay spread causes intersymbol interference due to frequency-selective fading and Doppler-spread results in additional signal distortion and temporal variation of the channel characteristics. However, if the channel characteristics can be estimated accurately, especially the Doppler-spread, many of these problems can be resolved. Doppler-spread occurs because the Doppler velocity may be different on each path. The paper describes a method for improved-reliability Doppler spread estimation of channels that experience frequency-selective fading. The method uses a wideband sounding signal to overcome the frequency selective fading and the estimation of the Doppler spread is obtained from this complex waveform using multi-layered optimization based on evolutionary algorithms. The method has a wide dynamic range and can estimate the Doppler shift of weak multipath components in the presence of strong multipath components. The accuracy of the Doppler-spread estimate is not limited by the short duration of sounding signal and the method performs well even in the presence of significant noise.

2:45–3:00 Break

3:00

2pSP6. Design of a symmetric multicarrier modulation for underwater acoustic communications. Yi Wang (Information and Coding Theory Lab., Univ. of Kiel, Kaiserstr. 2, D-24143 Kiel, Germany, yw@tf.uni-kiel.de)

Mobile underwater acoustic (UWA) channels always exhibit large delay spread and Doppler spread. Multicarrier modulation is an efficient technique for the communications over multipath channels. However, the orthogonality of the transmit signals, e.g., OFDM signals, will be destroyed by the Doppler spread and interference will be produced. Thus, an equalizer is required in the receiver. For this reason, we argue that multicarrier modulation with Gaussian pulse (symmetric in time and frequency domains) as the transmit filter is more suitable for the UWA communication systems than the OFDM. The key problem for the multicarrier modulation is to design/optimize the number of carriers and the parameter alpha in the Gaussian pulse which will greatly affect the system performance. In this paper, we investigate the optimization problem for four kinds of channels: AWGN channel, frequency-selective channel, time-selective channel and doubly selective channel. Channel information is assumed known in the transmitter by feedback. Analytical results are obtained and the numerical examples are given. To achieve the cost function, a 2D MMSE-DFE equalizer is used. Finally, we propose the balanced design for the practical engineering.

3:15

2pSP7. High-rate acoustic link for underwater video transmission using differential amplitude phase shift keying (DAPSK). Costas Pelekanakis, Milica Stojanovic (MIT, 292 Main St., Cambridge, MA 02139, gas@mit.edu), and Lee Freitag (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

A high bit-rate acoustic link for underwater video transmission is examined. Currently, video encoding standards allow video transmission at bit rates on the order of 64 kbps. To provide an acoustic transmission capability that meets this bit-rate requirement, we focus on the use of high-level bandwidth-efficient modulation methods. We consider the use of 16-, 32-, and 64-differential amplitude and phase shift keying (DAPSK) with varying number of amplitude levels. DAPSK does not require explicit carrier phase synchronization at receiver, but instead relies on differentially coherent detection. It thus represents an alternative to conventional quadrature amplitude modulation (QAM) methods, whose high-level constellations are sensitive to phase distortion induced by the channel. System performance is demonstrated experimentally, using 25 000 symbols/s at a carrier frequency of 75 kHz over a short vertical path. Excellent results were obtained, thus achieving bit rates as high as 150 kbps, which are sufficient for real-time transmission of compressed video.

3:30

2pSP8. Underwater acoustic communication experimentation: The acoustic communications and data storage (ACDS) system. M. McCord, J. Schindall, T. C. Yang, P. Gendron, and W.-B. Yang (Naval Res. Lab., 4555 Overlook Ave., Washington, DC 20375)

The Naval Research laboratory has developed a sea-going system, currently consisting of three acoustic communications and data storage (ACDS) buoys. The buoys can be deployed in the water column, moored to the ocean floor, or towed behind a ship to emulate an autonomous underwater vehicle (AUV). The system is intended for semi-real-time acoustic communications *in situ* and continuous recording of the raw acoustic data in the water for postexperiment analysis. The purpose is to study environmental issues affecting high data rate point-to-point acoustic communications as well as multiple-in-multiple-out acoustic networking. Each ACDS has an acoustic projector and an array of eight hydrophones. It is designed for two frequency bands: 2–5 kHz and 15–25 kHz. The system has been deployed in several oceans. Modulation signals used include binary and quadrature phase-shifted keying (BPSK/QPSK), frequency-hopped frequency-shifted keying (FH FSK), orthogonal frequency division multiplexing (OFDM) and code division multiple access (CDMA) signals. The acoustic modem is based on the Acoustic Modem Software developed by WHOI and improved by the Naval Underwater Warfare Center, Newport, RI. Acoustic networking signals are used to aid vehicle navigation. [Work supported by ONR.]

3:45

2pSP9. Doppler compensated underwater acoustic communication system. Anand Raj, Binu George, M. H. Supiya, James Kurian, and P. R. Saseendran Pillai (Dept. of Electron., Cochin Univ. of Sci. and Technol., Cochin 682 022, India)

Spread spectrum methods are used in communication systems to provide a low probability of intercept in hostile environments and multiple access capability in systems shared by many users as well as to provide high processing gain in channels where the transmitted signal is distorted by multipath effects. Such systems serve to be an effective tool for underwater telemetry environments, where multipath propagation effect and Doppler spreading is seen to be more predominant. This paper describes the implementation of a Doppler compensated underwater telemetry system based on CDMA technique. The system consists of multiple CDMA transmitters and a phase locked loop based carrier recoverable CDMA receiver. The effects of the Doppler shift can be compensated by the carrier recovery subsystem in the demodulator, based on PLL technique, which extracts the carrier frequency/phase and simultaneously demodulates the signal. The decision device in the receiver consists of a PN sequence generator as well as a bank of correlators, which are used to determine the data transmitted. The system simulation has been implemented in MATLAB. The advantage of this system is that multiple transmitting stations can transmit simultaneously to a central receiver, thereby increasing the system throughput.

4:00

2pSP10. Environmental and motion effects on orthogonal frequency division multiplexed signaling. Paul J. Gendron and T. C. Yang (4555 Overlook Ave. SW, Washington, DC 20375)

The use of tightly spaced orthogonal tones for communication through shallow water allows for diverse modulation strategies, from moderately efficient noncoherent on–off keying to fully coherent signaling. Critical to the effectiveness of orthogonal frequency division multiplexed (OFDM) schemes is the estimation of both source dilation, due to source–receiver relative motion, and the channel tranfer function. A dilation process estimator based on a spline model of dilation is presented and when coupled with decision directed channel estimation communication reliability is improved. The method is tested on OFDM signaling schemes at 18 kHz center frequency and 4 kHz bandwidth under diverse source motion scenarios and ranges during MREA(03) experiements off the north coast of Elba, Italy. Channel probing experiments characterize the environment in terms of coherence time, multi-path delay and angle spread. The quantitative effect of frequency selectivity and coherence time on the performance of OFDM signaling is addressed. Bit error rates are reported. [Work supported by ONR.]

TUESDAY AFTERNOON, 25 MAY 2004 VERSAILLES BALLROOM, 1:25 TO 3:30 P.M.

Session 2pUW

Underwater Acoustics: Ambient Noise

Anthony J. Eller, Chair
Science Applications International Corp., 1710 SAIC Drive, McLean, Virginia 22102

Chair's Introduction—1:25

Contributed Papers

1:30

2pUW1. Using ocean ambient noise for array element localization. Karim G. Sabra, Philippe Roux, W. A. Kuperman, W. S. Hodgkiss (Marine Physical Lab., Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093, ksabra@mpl.ucsd.edu), and Gerald L. D'Spain (Scripps Inst. of Oceanogr., UCSD, La Jolla, CA 92093-0704)

Time delays, associated with the direct and surface reflected arrivals, between the elements of a bottom hydrophone array can be extracted using ambient noise cross correlations [Sabra *et al.*, J. Acoust. Soc. Am. **114**, 2462 (2003)]. This is confirmed using long-time noise recordings that were collected in May 1995 near the S. California coast at an average depth of 21 m. The noise is mainly biological in the frequency range of 350–700 Hz [D'Spain *et al.*, J. Acoust. Soc. Am. **99**, 2453 (1996)]. These coherent wavefront arrivals across the array's aperture are used for array element localization.

1:45

2pUW2. Testing of ambient noise forecasting algorithms with LADC noise data. Anthony I. Eller (Sci. Applications Intl. Corp., 1710 Saic Dr., McLean, VA 22102, anthony.i.eller@saic.com), George E. Ioup, Juliette W. Ioup, and James P. Larue (Univ. of New Orleans, New Orleans, LA 70148)

Long-term, omnidirectional underwater acoustic noise was measured by three buoys in the northern Gulf of Mexico over a 36-day period during the summer of 2001 as part of the Littoral Acoustic Demonstration Center

project, phase I. This extensive data set is used to develop and evaluate algorithms for near-term temporal prediction of noise, based on immediate *in situ* data. Noise time series data in one-third-octave bands from 10 to 5000 Hz are processed to extract characterizing parameters such as mean level, variance, and temporal coherence across all frequency bands. The fluctuation spectrum is analyzed to determine how the noise variability is distributed with temporal period, and eigenvectors based on the fluctuation covariance matrix are used to identify underlying basic properties of noise variation. These tools are then used to predict future values of noise based on measurements at preceding times. [Research supported by ONR.]

2:00

2pUW3. Vertical directivity of ambient noise in the presence of internal waves. Richard B. Evans (Sci. Applications Intl. Corp., 23 Clara Dr., Mystic, CT 06355, richard.b.evans@saic.com)

Near surface sources of ambient noise, such as shipping and wind, excite only steeply traveling sound waves. After a sufficient range, the steepest are stripped away by loss mechanisms, such as bottom interaction. Consequently, the vertical directivity of surface generated noise is expected to be peaked about two (up- and down-going) angles with very little contribution near the horizontal. This expectation is not always met in measurements. A possible explanation, for the noise at the horizontal, is the random scattering of the up- and down-going waves by volume inhomogeneities caused by internal waves. The purpose of this paper is to demonstrate the validity of this hypothesis. The coupled power theory of Dozier and Tappert [J. Acoust. Soc. Am. **63**, 353–365 (1978)], in an updated form, is applied to the calculation of noise due to distant near surface sources. The vertical directivity of the resultant field is shown to demonstrate that internal wave scattering can produce noise at the horizontal. Rough surface scattering is expected to have a similar effect.

2:15

2pUW4. Ambient noise effects on towed-array search optimization. Donald R. DelBalzo and Erik R. Rike (Neptune Sci., Inc., 40201 Hwy. 190 E., Slidell, LA 70461)

Oceanographic and acoustic conditions in littoral environments are extremely complex and dynamic. Spatial and temporal variability of low-frequency signal and noise fields destroys the basic homogeneous assumptions associated with standard tactical search concepts, like simple ladder patterns. Genetic algorithms have recently been applied to the signal part of this problem [D. P. Kierstead and D. R. DelBalzo, ''A genetic algorithm approach for planning search paths in complicated environments,'' Military Operations Research Journal (March/April 2003)] to produce near-optimal, nonstandard search tracks that maximize probability of detection in uniform noise fields. In a complex environment, an optimal search plan will also depend on the spatial and temporal properties of AN. The dynamic ambient noise model (DANM) was used to produce low-frequency directional noise fields, based on discrete ship tracks. Then, optimal genetic search tracks were calculated for towed arrays, in the midst of spatial variability and temporal fluctuations of noise. The results show the importance of careful consideration of AN when designing optimum tactics. [Work sponsored by NAVSEA under the LCS project.]

2:30

2pUW5. An iso-deviant strategy for efficient ambient noise predictions with EAGLE. Erik R. Rike and Donald R. DelBalzo (Neptune Sci., Inc., 40201 Hwy. 190 E., Slidell, LA 70461, delbalzo@neptunesci.com)

Transmission loss (TL) computations in littoral areas require a dense spatial and azimuthal grid to achieve acceptable accuracy and detail, which is very slow. This problem of accuracy versus speed led to a new concept, OGRES (Objective Grid/Radials using Environmentally sensitive Selection), which produces sparse, irregular acoustic grids, with controlled accuracy. Recent work to further increase accuracy and efficiency with better metrics and interpolation led to EAGLE (Efficient Acoustic Gridder for Littoral Environments). On each iteration, EAGLE produces an acoustic field with approximately constant spatial uncertainty (hence, iso-deviance), yielding TL predictions with ever-increasing resolution and accuracy. This work adapts EAGLE to ambient noise computations. The Dynamic Ambient Noise Model (DANM) allows accurate, detailed estimation of the mean and variance of ambient noise in both space and time, but its TL computations are too slow for many applications. In the present work, a series of EAGLE acoustic field predictions was used by DANM (and compared to the dense full-grid solution) to determine the relationship between transmission loss uncertainty and noise-field uncertainty for a complex littoral area. An example is presented where approximately an order of magnitude efficiency improvement (over regular grids) is demonstrated. [Work sponsored by ONR under the LADC project.]

2:45

2pUW6. Ambient noise analysis of deep ocean measurements. Roy D. Gaul (Blue Sea Corp., Ste. 515, 14300 Cornerstone Village Dr., Houston, TX 77014), David P. Knobles (Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), and A. F. Wittenborn (TRACOR Sci. and Systems)

In October 1975 a measurement exercise designated CHURCH OPAL was done in the northeast Pacific Ocean to assess undersea acoustic noise and propagation phenomena. In 2003 the 10 days of deep ocean multiple hydrophone recordings during CHURCH OPAL were recovered and digitized. This paper presents results previously reported but unavailable for general distribution. The earlier work is augmented with more complete and detailed analyses using modern analytical techniques. Particular attention is given to statistical characterization of ambient noise within and beneath the deep sound channel in relation to distant shipping and local wind speed. [Work supported by ONR.]

3:00

2pUW7. Analysis on vertical directivity of shallow-water ambient noise in South China Sea. Ruey-Chang Wei (Inst. of Undersea Technol., Natl. Sun Yat-sen Univ., 70, Lien-hai Rd., Kaohsiung, Taiwan, rcwei@mail.nsysu.edu.tw), Hsiang-Chih Chan (Natl. Taiwan Univ., Taipei 804, Taiwan), and Po-Chang Lin (Natl. Sun Yat-sen Univ., Kaohsiung, Taiwan)

Beamforming is used to analyze the vertical directivity of the low-frequency ambient noise in shallow water. The data were measured by the vertical line array (VLA) of the Asian Seas International Acoustics Experiment (ASIAEX), from 3 May to 16 May 2001, and three frequencies, 200, 400, and 800 Hz, were studied. The results of the calculation show that the sound source of 200 and 400 Hz is from the horizontal direction and contributed by distant shipping, and the source of 800 Hz is mainly from high grazing angle that generated from surface noise. The well-known noise notch was also found in the study; the low-angle energy reflections from surface and bottom are the function of frequency, sound-speed profile, bathymetry, and the noise distribution. The time fluctuation of the noise notch is shown and correlated with oceanographic changes. The influences of internal wave and pass typhoon Cimaron during the experiment are significant too on vertical directivity of ambient noise.

3:15

2pUW8. Initial study of ambient noise modeling for 2001 ASIAEX SCS Experiment. Ruey-Chang Wei (Inst. of Undersea Technol., Natl. Sun Yat-sen Univ., 70, Lien-hai Rd., Kaohsiung, Taiwan, rcwei@mail.nsysu.edu.tw), Chi-Fang Chen (Natl. Taiwan Univ., Taipei, 804, Taiwan), and Hsiang-Chih Chan (Natl. Taiwan Univ., Taipei 804, Taiwan)

Although ambient noise has been an object of study for a long time, there is little modeling done for real ocean environment to reflect oceanographic changes. The numerical simulation in this study is based on PE code, with environmental and bathymetry data from the 2001 ASIAEX SCS experiment. Simple sources were distributed near sea surface in the model, and source levels in significant range were estimated by the wave height data of experiment. It is well known that strong internal waves

often occur in the South China Sea, which can cause significant change in water temperature distribution, thus the sound propagation. Moreover, the roughness of sea surface and bottom are also important factors in noise energy distribution. In this initial study, the fluctuation of noise energy would be evaluated by the varying spatial ocean environment. The computational noise levels would be compared with measurements on vertical line array in ASIAEX SCS experiment for adjusting the parameters and source levels.

TUESDAY AFTERNOON, 25 MAY 2004 CONFERENCE ROOM L, 2:00 TO 3:15 P.M.

Meeting of Accredited Standards Committee (ASC) S2 Mechanical Vibration and Shock

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meetings for:
ISO/TC 108, Mechanical Vibration and Shock
ISO/TC 108/SC 2 Measurement and evaluation of mechanical vibration and shock as applied to machines, vehicles and structures
ISO/TC 108/SC 3 Use and calibration of vibration and shock measuring instruments
ISO/TC 108/SC 5 Condition monitoring and diagnostics of machines
and
ISO/TC 108/SC 6 Vibration and shock generating systems

Accredited Standards Committee S2 on Mechanical Vibration and Shock. Working group chairs will report on the status of various shock and vibration standards currently under development. Consideration will be given to new standards that might be needed over the next few years. There will be a report on the interface of S2 activities with those of ISO/TC 108 and its subcommittees, including plans for future meetings of ISO/TC 108 and/or its Subcommittees. The Technical Advisory Groups for ISO/TC 108 and the Subcommittees listed above consists of members of S2 and other persons not necessarily members of those Committees. Open discussion of committee reports is encouraged.

Scope of S2: Standards, specifications, methods of measurement and test, and terminology in the 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.

TUESDAY AFTERNOON, 25 MAY 2004 CONFERENCE ROOM L, 3:30 TO 5:00 P.M.

Meeting of Accredited Standards Committee (ASC) S3 Bioacoustics

to be held jointly with the

ANSI-Accredited U.S. Technical Advisory Group (TAG) Meeting for:
ISO/TC 108/SC 4, Human Exposure to Mechanical Vibration and Shock

Accredited Standards Committee S3 on Bioacoustics. Working group chairs will report on the status of standards under development. Consideration will be given to new standards that might be needed over the next few years. There will be a report on the interface of S3 activities with those of ISO/TC 108/SC 4 Human exposure to mechanical vibration and shock, including plans for future meetings of ISO/TC 108/SC 4. The US Technical Advisory Group for TC 108/SC 4 consists of members of S3 and other persons not necessarily members of this Committee. Open discussion of committee reports is encouraged.

People interestred in attending the meeting of the TAGs for ISO/TC 43 Acoustics and IEC/TC 29 Electroacoustics, take note—those meetings will be held in conjunction with the Standards Plenary meeting at 8:00 a.m. on Tuesday, 25 May 2004.

Scope of S3: Standards, specifications, methods of measurements and test, and terminology in the fields of psychological and physiological acoustics, including aspects of general acoustics, shock, and vibration which pertain to biological safety, tolerance, and comfort.

Banquet

Dinner

6:30 p.m.–8:00 p.m.

Lawrence A. Crum and Anthony A. Atchley, Masters of Ceremonies

Program

Welcome Remarks by Ilene J. Busch-Vishniac, President

Presentation of Certificates to New Fellows

Abeer Alwan	Jeffrey A. Nystuen
David H. Chambers	Andrew J. Oxenham
Robin O. Cleveland	Christopher J. Plack
Li Deng	Daniel Rouseff
Alain de Cheveigne	Carrick Talmadge
Gary J. Heald	D. Keith Wilson
Jian Kang	Eric J. Woods

Viewing of Special Film made for the
75th Anniversary Celebration

Closing Remarks

Session 4aAA

Architectural Acoustics: Acoustics of Virtual Environments

Elizabeth A. Cohen, Cochair
Cohen Acoustical, Inc., California 90004

Mendel Kleiner, Cochair
School of Architecture, Rensselaer Polytechnic Institute, 110 8th Street, Troy, New York 12180-3590

Chair's Introduction—8:30

Invited Papers

8:35

4aAA1. Application-driven design of auralization systems. Durand R. Begault (NASA Ames Res. Ctr., MS 262-2, Moffett Field, CA 94035-1000, Durand.R.Begault@nasa.gov)

Acoustical environmental simulation (auralization) involves rendering a binaural signal that articulates simultaneously the positional information of a sound source and the source's acoustic and vibratory interaction with its environmental context. This places significant challenges on both the modeling and rendering components of such a system. Overcoming real-time limitations of processors can be accomplished by implementing rendering limits based on auditory threshold data for both early reflections and late reverberant energy. However, final assessment of system quality is dependent on the specific task or goal of the simulation. The use of auralization in accomplishing a telerobotic task requires head tracking and low latency, while environmental cues such as obstacles with acoustic reflections can be simplified and exaggerated beyond veridical representation. Contrasting this is the task of assessing speech intelligibility of an emergency public address system, or the quality of a space for musical performance, requiring a more accurate level of simulation of acoustical materials and sound-source characteristics. Modeling accuracy is complicated by such factors as the level of variance in the absorptive and diffusive properties of materials in terms of their real-world application. Assessment of auralization quality must therefore involve a best estimate of significant factors for a particular application.

8:55

4aAA2. The importance of perceived self-motion in experiencing convincing virtual acoustic rendering. William L. Martens (Faculty of Music, McGill Univ., Montreal, QC H3A 1E3, Canada, wlm@music.mcgill.ca)

One factor that makes a virtual acoustic environment more convincing is the capacity to present information that enables a listener to answer questions about the geometry of the acoustical space in which they are immersed. This paper reviews the results of several studies in which virtual sound sources were presented to listeners via a loudspeaker array located in an anechoic chamber, and asked about details such as the location of a missing wall in a simulated rectangular room, and whether or not a reduction in the loudness of a virtual source was due to the interposition of a virtual wall between source and listener. When a listener was unaware of changes in the virtual listening position, such questions were difficult to answer; however, when listeners were informed of their path through the environment via computer graphic animation, performance improved significantly. The results of these perceptual experiments suggest that perceived self-motion may be an important determinant of how convincing virtual acoustic rendering may be. The identity of the sound source itself was also found to have a substantial impact, since listeners reported self-motion more often when presented with sound sources that were judged less likely to be in motion.

Contributed Papers

9:15

4aAA3. Active field control for piano room. Atsuko Ito and Yasushi Shimizu (Adv. System Development Ctr., Yamaha Corp., 10-1 Nakazawa-cho, Hamamatsu 430-8650, Japan, aito@beat.yamaha.co.jp)

Active field control (AFC) is an innovative system that employs electroacoustic technologies enabling ''natural changes'' to major auditory impressions, such as reverberance, loudness, and spaciousness. This talk presents typical design concept of AFC application for a piano room with a low ceiling in architecture and with a miniconcert in program. The aim is to accommodate a condition appropriate for both a solo piano performance and a miniconcert, regardless of the low ceiling height and significant increase of absorption by the audience in a miniconcert. This can be realized by using an AFC system, which converts the room into different acoustical conditions as follows. One is that AFC alters acoustics of the designed space, and makes it sound as if it has a higher ceiling, which leads to a longer reverberation. The other is negative absorption to compensate added absorption with the audience for a miniconcert. In the presentation, these technologies will be presented, along with the acoustical data.

9:30

4aAA4. Assessing the accuracy of auralizations computed using a hybrid geometrical-acoustics and wave-acoustics method. Jason E. Summers, Kengo Takahashi, Yasushi Shimizu, and Takashi Yamakawa (Adv. Syst. Dev. Ctr., Yamaha Corp. of Japan, 10-1 Nakazawa-cho, Hamamatsu 430-8650, Japan, jesummers@alumni.kenyon.edu)

When based on geometrical acoustics, computational models used for auralization of auditorium sound fields are physically inaccurate at low frequencies. To increase accuracy while keeping computation tractable, hybrid methods using computational wave acoustics at low frequencies have been proposed and implemented in small enclosures such as simplified models of car cabins [Granier *et al.*, J. Audio Eng. Soc. **44**, 835–849

(1996)]. The present work extends such an approach to an actual 2400-m^3 auditorium using the boundary-element method for frequencies below 100 Hz. The effect of including wave-acoustics at low frequencies is assessed by comparing the predictions of the hybrid model with those of the geometrical-acoustics model and comparing both with measurements. Conventional room-acoustical metrics are used together with new methods based on two-dimensional distance measures applied to time-frequency representations of impulse responses. Despite *in situ* measurements of boundary impedance, uncertainties in input parameters limit the accuracy of the computed results at low frequencies. However, aural perception ultimately defines the required accuracy of computational models. An algorithmic method for making such evaluations is proposed based on correlating listening-test results with distance measures between time-frequency representations derived from auditory models of the ear-brain system. Preliminary results are presented.

9:45

4aAA5. A binaural Web-based tour of the acoustics of Troy Music Hall. Rendell R. Torres, James Cooney, and Yasushi Shimizu (Prog. in Architectural Acoust., Rensselaer Polytech. Inst., Greene Bldg., 110 8th St., Troy, NY 12180-3590, rrtorres@rpi.edu)

For classical music to become more widely enjoyed, it must sound exciting. We hypothesize that if people could hear examples of truly exciting acoustics, classical music would be perceived less as a rarefied delicacy and more as a viscerally engaging listening experience. The Troy Savings Bank Music Hall in Troy, New York, is a legendary 1200-seat concert hall famous for its acoustics. Such landmarks are commonly documented architecturally but with few attempts to document their acoustics in a way that it is listenable. Thus, the goal is to capture and sonically disseminate the hall's acoustics through a Web-based acoustical tour, where one can click on various seats to hear binaural auralizations of different instruments and see corresponding views of the stage. The hope is that these auralizations will not only sonically document the acoustics of the hall but also tantalize even geographically distant listeners with binaural samples of how exciting music can be in excellent acoustics. The fun and challenges of devising (let alone standardizing) such an auralization-based system of documentation will be discussed, and a demonstration given. This process can be applied to other historically and acoustically significant spaces. [Work supported by the National Endowment for the Arts.]

10:00–10:15 Break

10:15

4aAA6. Virtual room modeling using hybrid digital waveguide mesh techniques. Mark J. Beeson and Damian T. Murphy (Media Eng. Group, Dept. of Electron., Univ. of York, Heslington, York, UK)

Digital waveguide mesh (WGM) models have been shown to be a viable method of obtaining accurate room impulse responses (RIRs) for a virtual space. However, the large memory and long processing requirements of these models have restricted their use for large rooms and for those with nontrivial geometric features. The development and inclusion of a KW-pipe interface allows the interconnection of finite difference and wave-based mesh implementations of a WGM model. The former is efficient for the main body of the mesh, with the latter, more processor intensive method allowing more accurate simulation at the boundaries of the modeled space. The resultant hybrid model removes some of the above computational constraints, allowing larger rooms and more complex geometries to be modeled. Model visualizations and RIR data are presented, demonstrating the correct operation of the KW-pipe interface. RIR data from two differing mesh topologies show that nontrivial geometries can be successfully modeled using this technique, with considerable computational savings on purely wave-based WGMs. Comparisons with RIRs obtained through more traditional ray-tracing and image-source techniques show favorable results and demonstrate the complex wave phenomena that are inherent in WGM models. [Work supported by EPSRC.]

10:30

4aAA7. Acoustical analysis and multiple source auralizations of charismatic worship spaces. Richard W. Lee (Rensselaer Polytech. Inst., Prog. in Architectural Acoust., 110 8th St., Troy, NY 12180-3590)

Because of the spontaneity and high level of call and response, many charismatic churches have verbal and musical communication problems that stem from highly reverberant sound fields, poor speech intelligibility, and muddy music. This research looks at the subjective dimensions of room acoustics perception that affect a charismatic worship space, which is summarized using the acronym RISCS (reverberation, intimacy, strength, coloration, and spaciousness). The method of research is to obtain acoustical measurements for three worship spaces in order to analyze the objective parameters associated with the RISCS subjective dimensions. For the same spaces, binaural room impulse response (BRIR) measurements are done for different receiver positions in order to create an auralization for each position. The subjective descriptors of RISCS are analyzed through the use of listening tests of the three auralized spaces. The results from the measurements and listening tests are analyzed to determine if listeners' perceptions correlate with the objective parameter results, the appropriateness of the subjective parameters for the use of the space, and which parameters seem to take precedent. A comparison of the multi-source auralization to a conventional single-source auralization was done with the mixed down version of the synchronized multi-track anechoic signals.

10:45

4aAA8. Sensitivity of room acoustic parameters to changes in scattering coefficients. Jonathan Rathsam and Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, jrathsam@mail.unomaha.edu)

This project uses the room acoustics computer modeling program, ODEON, to investigate the sensitivity of room acoustic parameters to changes in scattering coefficients. Particularly, the study is interested in determining if the results from certain room models are more sensitive to scattering coefficients than from other models, due to their geometry or absorption characteristics. If so, how can one quantify a model's susceptibility to being sensitive to scattering? Various models of three real spaces in Omaha, Nebraska are tested. The predicted reverberation, clarity, and spaciousness parameters are compared at various receiver locations, while the scattering coefficient of all surfaces is varied from 0 to 0.1, 0.3, 0.5, and 0.8. The resulting data are analyzed by frequency according to the (1) average absorption of the room; (2) magnitude variation of absorption within the room; (3) spatial distribution of absorption within the room; and (4) level of model detail. Initial results indicate that parameters studied may show more sensitivity to scattering coefficients in models that have a wider range of absorption values, more disparate distribution of absorption, and lower detail level. Various schemes that include these aspects are proposed for computing a model's sensitivity to changes in scattering.

11:00

4aAA9. Investigation of the optimum acoustical conditions for speech using auralization. Wonyoung Yang and Murray Hodgson (School of Occupational and Environ. Hygiene, Univ. of BC, 2206 East Mall, Vancouver, BC V6T 1Z3, Canada, wyang@interchange.ubc.ca)

Speech intelligibility is mainly affected by reverberation and by signal-to-noise level difference, the difference between the speech-signal and background-noise levels at a receiver. An important question for the design of rooms for speech (e.g., classrooms) is, what are the optimal values of these factors? This question has been studied experimentally and theoretically. Experimental studies found zero optimal reverberation time, but theoretical predictions found nonzero reverberation times. These contradictory results are partly caused by the different ways of accounting for background noise. Background noise sources and their locations inside the room are the most detrimental factors in speech intelligibility. However,

noise levels also interact with reverberation in rooms. In this project, two major room-acoustical factors for speech intelligibility were controlled using speech and noise sources of known relative output levels located in a virtual room with known reverberation. Speech intelligibility test signals were played in the virtual room and auralized for listeners. The Modified Rhyme Test (MRT) and babble noise were used to measure subjective speech intelligibility quality. Optimal reverberation times, and the optimal values of other speech intelligibility metrics, for normal-hearing people and for hard-of-hearing people, were identified and compared.

11:15

4aAA10. An evaluation of differences due to changing source directivity in room acoustic computer modeling. Michelle C. Vigeant and Lily M. Wang (Architectural Eng. Prog., Univ. of Nebraska–Lincoln, Peter Kiewit Inst., 1110 S. 67th St., Omaha, NE 68182-0681, mvigeant@unlnotes.unl.edu)

This project examines the effects of changing source directivity in room acoustic computer models on objective parameters and subjective perception. Acoustic parameters and auralizations calculated from omnidirectional versus directional sources were compared. Three realistic directional sources were used, measured in a limited number of octave bands from a piano, singing voice, and violin. A highly directional source that beams only within a sixteenth-tant of a sphere was also tested. Objectively, there were differences of 5% or more in reverberation time (RT) between the realistic directional and omnidirectional sources. Between the beaming directional and omnidirectional sources, differences in clarity were close to the just-noticeable-difference (jnd) criterion of 1 dB. Subjectively, participants had great difficulty distinguishing between the realistic and omnidirectional sources; very few could discern the differences in RTs. However, a larger percentage (32% vs 20%) could differentiate between the beaming and omnidirectional sources, as well as the respective differences in clarity. Further studies of the objective results from different beaming sources have been pursued. The direction of the beaming source in the room is changed, as well as the beamwidth. The objective results are analyzed to determine if differences fall within the jnd of sound-pressure level, RT, and clarity.

11:30

4aAA11. Physical and subjective study of randomized sound scattering through geometrical acoustic simulation. Weifang Wang (Shen Milsom & Wilke, 417 Fifth Ave., New York, NY 10016) and Rendell R. Torres (Rensselaer Polytechnic Inst., Troy, NY 12180)

Sound scattering is one of the most important phenomena to consider for accurate and fine detailed calculation in both room acoustic design and simulation. Up to now much research work has been done on sound scattering measurement and design and calculation, but how it affects or is relevant to our perception still remains unknown to us. The aim of this research is just trying to find out how randomized sound scattering changes affect our perception and what kind of problems still exist with the current research method based on geometrical acoustic simulation. This research has been carried out with psychoacoustic experiments and advanced time-frequency analysis based on geometric acoustic simulation and binaural signal synthesis. Statistical conclusion on effects of sound scattering change and corresponding visual representation would be given thereafter.

11:45

4aAA12. On the simulation of seat-dip effect using geometrical acoustics software. Ettore Cirillo and Francesco Martellotta (Dip. Fisica Tecnica, Politecnico di Bari, via Orabona 4, 70125 Bari, Italy, f.martellotta@poliba.it)

A group of Italian churches was surveyed in order to measure the most important acoustic parameters according to ISO 3382 Standard. Computer models of the same churches were made using acoustic simulation software. Absorption coefficients found in the literature were used and later calibrated to match predicted and measured T30 values. The results of the simulations were compared with the observed values, showing some discrepancies at mid-low frequencies. This discrepancy appeared to be due to a lack of direct sound in the measured responses, particularly at the 250-Hz and 500-Hz octave bands, indicating the probable presence of a seat-dip effect caused by the wooden pews. Since the acoustic simulation software provided the possibility to use acoustically semitransparent planes, this feature was used to simulate the effect of selective absorption of the direct sound due to the seat-dip effect. The comparison between measured acoustic parameters and those predicted including the simulation of the seat-dip effect showed that an improvement in the prediction accuracy can be achieved. Different configurations were tested in order to define the optimal placing of the semitransparent plane, and a criterion to choose the transparency coefficients is finally proposed.

THURSDAY MORNING, 27 MAY 2004 LIBERTY 3, 8:00 A.M. TO 12:05 P.M.

Session 4aAB

Animal Bioacoustics: Natural Acoustic Behavior of Animals: Session in Memory of Donald R. Griffin III

Andrea M. Simmons, Chair
Department of Psychology, Brown University, Providence, Rhode Island 02912

Invited Paper

8:00

4aAB1. Acoustic behavior of echolocating bats in complex environments. Cynthia Moss, Kaushik Ghose, Marianne Jensen (Dept. of Psych., Inst. for Systems Res., Univ. of Maryland, College Park, MD 20742, cmoss@psyc.umd.edu), and Annemarie Surlykke (Univ. of Southern Denmark, Odense, Denmark)

The echolocating bat controls the direction of its sonar beam, just as visually dominant animals control the movement of their eyes to foveate targets of interest. The sonar beam aim of the echolocating bat can therefore serve as an index of the animal's attention to objects in the environment. Until recently, spatial attention has not been studied in the context of echolocation, perhaps due to the difficulty in obtaining an objective measure. Here, we describe measurements of the bat's sonar beam aim, serving as an index of

acoustic gaze and attention to objects, in tasks that require localization of obstacles and insect prey. Measurements of the bat's sonar beam aim are taken from microphone array recordings of vocal signals produced by a free-flying bat under experimentally controlled conditions. In some situations, the animal relies on spatial memory over reflected sounds, perhaps because its perceptual system cannot easily organize cascades of echoes from obstacles and prey. This highlights the complexity of the bat's orientation behavior, which can alternate between active sensing and spatial memory systems. The bat's use of spatial memory for orientation also will be addressed in this talk. [Work supported by NSF-IBN-0111973 and the Danish Research Council.]

Contributed Papers

8:20

4aAB2. Dolphin and bat sonar: Convergence, divergence, or parallelism. Darlene R. Ketten (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543 and Harvard Med. School, Boston, MA 02114, dketten@whoi.edu), James Simmons (Brown Univ., Providence, RI 02912), Allyn E. Hubbard, and David A. Mountain (Boston Univ., Boston, MA)

During the explosive period of mammalian radiation, two groups emerged with highly effective biosonar systems, bats and toothed whales. In the intervening 50 million years, these groups evolutionarily honed their hearing for operation in radically different media. This paper addresses what functional aspects the media influenced in the biosonar receptors of bats versus dolphins by comparing the auditory peripheries of these groups. Data were obtained using thin-section microscopy, CT imaging, and inner-ear models. Inner-ear anatomy is fundamentally similar in these animals, although differences exist in both neural density and distribution in each group. Specialist ears are present in both groups, suggesting at least one odontocete species has cochlear specializations consistent with CF-FM bats, including specialized basilar-membrane regions and high-frequency neural foveal areas. Cochlear specializations in both groups are primarily linked to peak spectra of sonar signals, may expand frequency representation, and may enhance tuning in adjacent ear segments by generating standing wave phenomena. Most differences, such as the soft-tissue external ear analogs in odontocetes, are clearly media driven. Other differences among species within each group are correlated with signal type or habitat complexity. [Work supported by Mellon Foundation; Seaver Institute; ONR.]

8:35

4aAB3. Biosonar signal processing of bats during flight observed by a telemetry microphone on the head. Hiroshi Riquimaroux, Yoshiaki Watanabe (Fac. Eng., Doshisha Univ., Kyotanabe 610-0321, Japan, hrikimar@mail.doshisha.ac.jp), and Liang-Kong Lin (Tunghai Univ., Taichung, Taiwan)

The purpose of the present study was to examine the daily consistency of constant frequency component (CF) of CF-FM bats and to examine biosonar sound characteristics of leaf-nosed bats (*Hipposiderous terasensis*) during flight by a telemetry microphone system (telemike) mounted on the bat's head. Experiments were done within a steel-walled chamber. When at rest, the fundamental frequency of the bat was about 35 kHz. The second harmonic (70 kHz) was the most intense. The second harmonic of CF component (CF2) at rest and during flight varied significantly over several months. A systematical change in CF2 frequency during flight was observed through the telemike. Doppler-shift compensation in CF2 frequency, amplitude compensation, and pulse emission rate just before landing were confirmed. We will discuss time-sharing dual signal processing found during flight. [Research supported by Special Coordination Funds, and a grant to RCAST at Doshisha University from the Ministry of Education, Culture, Sports, Science and Technology of Japan.]

8:50

4aAB4. A portable system for marine mammal auditory-evoked potential measurements. James J. Finneran (U.S. Navy Marine Mammal Prog., Space and Naval Warfare Systems Ctr., San Diego, Code 2351, 53560 Hull St., San Diego, CA 92152, james.finneran@navy.mil) and Dorian S. Houser (Biomimetica, La Mesa, CA 91942)

Limitations to behavioral measures of hearing sensitivity in marine mammals include the time and expense typically required to train subjects. These limitations have resulted in limited subjects and lingering questions regarding intraspecific variability. An alternative to behavioral methods is the electrophysiological method, where passive electrodes are used to measure auditory-evoked potentials (AEPs) generated by the brain in response to sound stimuli. Marine mammal AEP measurements have been limited by the complexity of the technique and the limited applicability of commercially available AEP systems. In this paper, a portable, laptop computer-based system for marine mammal AEP measurements will be presented. The system features commercial off-the-shelf components, including a data acquisition PC card, biopotential amplifier, and programmable attenuator. The system is housed in a rugged, shock-resistant case. Custom software is used to present sound stimuli, record evoked responses, and analyze the resulting data. The system has been used to measure auditory brainstem responses to clicks and tone pips and envelope following responses to amplitude-modulated tones in bottlenose dolphins. Preliminary data obtained with the system will be presented and compared to behavioral hearing measures. [Work supported by the ILIR at SPAWARSYSCEN-SD and the ONR.]

9:05

4aAB5. Ecological echoes observed by moving biomimetic sonar characterize objects. Roman Kuc (Dept. of Elec. Eng., Yale Univ., New Haven, CT 06520-8267)

This paper examines echoes from *in situ* foliage, similar to those observed by flying bats. A moving sonar converts echoes into spike sequences and applies neural-computational methods to estimate passing range and travel speed. Retro-reflectors and favorably oriented planar patches generate strong echoes (SEs), identified by spike firing rate. Linear sonar trajectories cause SEs to form hyperbolic patterns, termed *glints*, specified by passing range and travel speed. Passing-range specific detectors compare successive SE times with values in a table and tally coincidences. A glint terminates after a sufficient number of coincidences are tallied and two consecutive mismatches occur in the maximum-count detector. SE arrival jitter necessitates a coincidence window. Short windows identify individual glints while long windows generalize extended objects. SEs from distant objects exhibit almost constant incremental delays, used to estimate sonar travel speed, necessary for robust glint detection. Passing-range estimates may explain how bats can fly through small openings without collision.

9:20

4aAB6. How bats' ears probe space: A numerical analysis of pinna shapes. Rolf Müller, John C. T. Hallam (Maersk Inst., Univ. of Southern Denmark, DK-5230 Odense M, Denmark, rolfm@mip.sdu.dk), Herbert Peremans (Univ. of Antwerp, BE-2000 Antwerp, Belgium), Alexander Streicher, and Reinhard Lerch (Univ. of Erlangen-Nürnberg, D-91052 Erlangen, Germany)

Like any antenna, bat pinnae have directivity patterns which impose a frequency-dependent sensitivity weighting on possible locations of sonar targets. Here, numerical predictions of such directivities are presented. The predictions are based on finite-element meshes generated from computer-tomographic cross-section images of pinna samples obtained from several bat species. A time-domain finite-element model was combined with a near-field to far-field transformation to yield high-resolution directivity estimates. The data obtained enable the introduction of new visualization approaches to the study of bat biosonar. These techniques portray the spatial and frequency dimension of the patterns, facilitate quantitative analysis, and reveal functional properties relevant to specific sensing tasks, like scanning the surface of an extended target. The computer graphics representations of the pinna shapes (voxel arrays, volume and surface meshes) can be readily manipulated to shed light onto how the

observed properties arise. For example, appendages have been removed by ''Boolean surgery'' or altered in orientation or shape by standard computer graphics transformations. The results of these manipulations are compared with respect to wave-field amplitudes as well as with respect to directivity patterns. [Work supported by the European Union (CIRCE Project, IST-2001-35144).]

9:35

4aAB7. Digital neuromorphic processing for a simplified algorithm of ultrasonic reception. Lin Qiang and Chris Clarke (Dept. of Elec. & Electron. Eng., Univ. of Bath, Bath BA2 7AY, UK, L.Qiang@bath.ac.uk)

Previously, most mammalian auditory systems research has concentrated on human sensory perception whose frequencies are lower than 20 kHz. The implementations almost always used analog VLSI design. Due to the complexity of the model, it is difficult to implement these algorithms using current digital technology. This paper introduces a simplified model of biosonic reception system in bats and its implementation in the ''Chiroptera Inspired Robotic CEphaloid'' (CIRCE) project. This model consists of bandpass filters, a half-wave rectifier, low-pass filters, automatic gain control, and spike generation with thresholds. Due to the real-time requirements of the system, the system employs Butterworth filters and advanced field programmable gate array (FPGA) architectures to provide a viable solution. The ultrasonic signal processing is implemented on a Xilinx FPGA Virtex II device in real time. In the system, 12-bit input echo signals from receivers are sampled at 1 M samples per second for a signal frequency range from 20 to 200 kHz. The system performs a 704-channel per ear auditory pipeline operating in real time. The output of the system is a coded time series of threshold crossing points. Comparing hardware implementation with fixed-point software, the system shows significant performance gains with no loss of accuracy.

9:50

4aAB8. Off-axis signal processing of cetacean biosonar. Walter M. X. Zimmer (NATO Undersea Res. Ctr., Viale San Bartolomeo 400, I-19138 La Spezia, Italy, walter@saclantc.nato.int)

Echolocation or biosonar plays a fundamental role for odontocetes to probe their environment, and their characteristics have been studied extensively for over 40 years. In summary, cetacean biosonar can be modeled as broadband transient-like signals radiating from a finite piston or aperture. The resulting sonar beam is directional, with a directivity index exceeding 25 dB for some species. The estimation of relevant sonar parameters is usually obtained from animals that are kept within a controlled and well-instrumented environment. This paper shows how measurements of opportunity from free-ranging odontocetes may be used to obtain their relevant biosonar parameters, in particular directivity index and source level. Frequently these measurements are made with a single hydrophone that is sufficiently deep so that surface-reflected echoes separate from the direct arrival of the echolocation clicks. Also, the received signal often fades in and out as the sonar beam of the scanning animal crosses the hydrophone. The presented technique exploits this scanning and the observation that broadband signals from a finite aperture will appear distorted when recorded off the acoustic axis, as the transfer function of the aperture modifies the spectrum of the transmitted signal.

10:05–10:20 Break

10:20

4aAB9. The whistles of bottlenose dolphins (*Tursiops truncatus*) from the Gulf of Mexico. Carmen Bazua-Duran (Lab. Acustica Aplicada y Vibraciones, CCADET, UNAM, Cd. Universitaria, 04510 Mexico DF, Mexico, bazua@servidor.unam.mx)

This work presents the description and geographic comparison of whistles from bottlenose dolphins recorded in three coastal areas in the northwestern Gulf of Mexico (Galveston and Corpus Christi Bays and Madre Lagoon), one oceanic area in the northern Gulf, and one coastal area in the southern Gulf (Terminos Lagoon). The 1499 whistle contours analyzed were categorized into 289 whistle types, of which 120 types were unique to a specific area. From the remaining 169 types, more types were common between Madre and Corpus Christi than between Galveston and either Corpus Christi or Madre, results in agreement with the dolphin mixing patterns between these three areas. Whistles from the oceanic area were more similar to those of Madre, suggesting that contact between coastal and oceanic dolphins in the northwestern part of the Gulf may be through Madre Lagoon. Terminos and Galveston whistles, the areas furthest apart, were very similar, indicating that contact between dolphin groups may not be the only parameter determining whistle repertoire similarities. Dolphin whistle similarities may also depend on comparable habitat use and population structure. The new signal type curve shows that more than 250 whistles are needed for each area in order to adequately describe the whistle repertoire.

10:35

4aAB10. Free-ranging finless porpoises acoustically inspect their frontal area in advance. Tomonari Akamatsu (Natl. Res. Inst. of Fisheries Eng., Ebidai, Hasaki, Kashima, Ibaraki 314-0421, Japan, akamatsu@affrc.go.jp), Ding Wang, Kexiong Wang (Chinese Acad. of Sci., 430072, PROC), and Yasuhiko Naito (Natl. Inst. of Polar Res., Tokyo 173-8515, Japan)

Echolocation events, interpulse intervals, and swimming speeds of nine free-ranging finless porpoises in an oxbow of the Yangtze River, China were recorded by datalogger systems attached on the animals. Over 120 h of successful recording indicated that the finless porpoises acoustically inspected their frontal area in advance before swimming silently. The acoustical sensing distance estimated by the interpulse interval was significantly larger than the swimming distance without echolocation beforehand. Terminal phase which was already known in the echolocation behavior of bats could be found in free-ranging finless porpoises. The terminal phase is the decreasing interpulse intervals in an echolocation pulse train that are observed just before the prey capture. During the terminal phase of finless porpoises, linearly decreased interpulse intervals were recognized. In the mean time, the swimming distance and the change of the sensing distance were closely correlated with each other. This suggests that the finless porpoise knew precisely the distance to the approaching target in the time scale of subsecond order. Acoustical sensing effort was considered to be controlled appropriately by free-ranging finless porpoises to obtain underwater information they need. [Research supported by Promotion of Basic Research Activities for Innovative Biosciences, Bio-oriented Technology Research Advancement Institution, Japan.]

10:50

4aAB11. Passive tracking of multiple diving sperm whales using single hydrophones at two mobile locations. Rajendar Bahl (Univ. of Tokyo, Underwater Technol. Res. Ctr., Inst. of Industrial Sci., 4-6-1 Komaba, Meguro, Tokyo 153-8505, Japan, rbahl@iis.u-tokyo.ac.jp), Takeshi Nakatani, Tamaki Ura (Univ. of Tokyo, Tokyo, Japan), Junichi Kojima (KDDI R&D Labs., Inc., Saitama, Japan), Tetsuo Fukuchi (System Giken Corp., Kanagawa, Japan), Masao Sakata, Yoshiaki Nose (Univ. of Tokyo, Tokyo, Japan), Junya Ura (Waseda Univ., Tokyo, Japan), Kyoichi Mori (Ogasawara Whale Watching Assn., Tokyo 100-2101, Japan), Harumi Sugimatsu (Univ. of Tokyo, Tokyo 153-8505, Japan), Masao Yanagisawa (Waseda Univ., Tokyo, Japan), and Hideyuki Suzuki (Univ. of Tokyo, Tokyo, Japan)

A simple method using only two hydrophones, each loosely deployed from separate mobile platforms, has been developed to simultaneously track several vocalizing sperm whales. The separation distance of several hundred meters between the hydrophones implies that a particular whale could present vastly different beam orientations towards them, thus precluding use of any specific relationship between the click signal levels at the two locations for the purpose of whale localization. This method utilizes time-of-arrival of the direct clicks and their surface reflections and matches them at the two hydrophone locations. Whales are segregated on the basis of the set of observed time delays. Click parameters such as interpulse interval and average click frequency are proposed to be used as secondary data for the purpose of track refinement. Depth profile of the

whales is obtained independent of the separation distance between the two hydrophones. However, knowledge of the separation distance between the hydrophones provides the 3D coordinates of the whales within a left–right ambiguity. Whale tracks have been obtained using actual data collected off Ogasawara Islands in Japan. These results demonstrate the utility of this method for studying the bioacoustics and behavior of deep diving sperm whales.

11:05

4aAB12. The "gunshot" sound produced by male North Atlantic right whales and its potential function in reproductive advertisement. Susan E. Parks (Cornell Bioacoustics Res. Prog., 159 Sapsucker Woods Rd., Ithaca, NY 14850, sep6@cornell.edu), Philip Hamilton, Scott D. Kraus (New England Aquarium, Boston, MA 02110), and Peter L. Tyack (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

North Atlantic right whales (*Eubalaena glacialis*) commonly use sound to mediate social interactions between individuals. Surface active groups (SAGs) are the most commonly observed social interaction on the summer feeding grounds. These groups are typically composed of an adult female with two or more males engaged in social behavior at the surface. Several distinct types of sounds have been recorded from these groups. One sound commonly recorded from these groups is a brief broadband sound, referred to as a gunshot sound because it sounds like a rifle being fired. This sound has been recorded in the Bay of Fundy, Canada from both lone whales ($N=9$) and social SAGs ($N=49$). Those lone whales producing gunshot sounds whose sex could be determined ($N=9$) were all mature males. In surface active groups, the rate of production of gunshot sounds was weakly correlated with the total number of males present in the group. Given the behavioral contexts of gunshot sound production by male whales, gunshots probably function in a reproductive context as an agonistic signal directed toward other males, an advertisement signal to attract females, or a combination of the two functions.

11:20

4aAB13. Influence of ambient noise on the use of sound by marine animals. Douglas H. Cato (Defence Sci. & Tech Org., and Univ. of Sydney Inst. of Marine Sci., P.O. Box 44, Pyrmont, NSW 2009, Australia, doug.cato@dsto.defence.gov.au)

Ambient noise provides the basic limitation on the use of sound by animals since signals of interest must be detected against the noise. In the ocean, ambient noise is very variable, both temporally and spatially. Temporal variation is more than 30 dB, with 20 dB being common and occurring over time scales of order 1 day, causing substantial variation in audible distances of sources (order a factor of 10). This paper examines how ambient noise affects the use of sound by animals and how it may have shaped aspects in development acoustic function. It draws on a large amount of data from the Australian region which is relatively free of anthropogenic noise, and contains a wide range of biological activity. Characteristics of ambient noise are compared with those of some marine animal vocalizations to assess whether these have any optimization against the noise. Although low acoustic absorption loss in water provides the potential for sound to be used over large distances, the variability of ambient noise and propagation suggests that high source levels of marine animals are needed to allow efficient use of sound over moderate distances most of the time, rather than large distances under exceptional conditions.

11:35

4aAB14. The impact of observer dynamics on sonar perception. Richard J. Rikoski (NSWC Panama City, Code R11—Robotic Technologies Branch, 110 Vernon Ave., Panama City, FL 32407, RikoskiRJ@ncsc.navy.mil)

The minimum information necessary for a moving observer to track objects depends on body dynamics, sensor design, sensor configuration, and environmental perturbations. This paper will present governing equations relating sonar measurements to 6 DOF dynamics, and eight nondimensional parameters describing the transitions between various regimes. Using nondimensional analysis, it will be shown that to minimize second-order effects, an observer should increase its ping rate as it approaches a target. Bats have been observed to increase their ping rate as they approach a target [Altringham, *Bats: Biology and Behavior* (1996)]. Alternatively, an observer might use a target's angular rate to predict second-order range effects. This requires the angular rate to become observable prior to second-order range effects becoming observable. This yields a design constraint relating velocity, ping rate, and aperture, but not wavelength. Based on this constraint, it appears some marine mammals could track objects in the second-order regime [Au, *The Sonar of Dolphins* (1993)]. [Work supported by ONR.]

11:50

4aAB15. Model-based localization and tracking of marine mammals. Ahmad T. Abawi, Michael B. Porter, Martin Siderius (Ctr. for Ocean Res., Sci. Applications Intl. Corp., 10260 Campus Point Dr., San Diego, CA 92121), John Hildebrand, and Sean Wiggins (Scripps Inst. of Oceanogr., LaJolla, CA)

Data from the August 2003 experiment conducted by the Scripps Institution of Oceanography in the Southern California Offshore Range (SCORE) are used to localize and track marine mammals. SCORE is a naval training area near the island of San Clemente located in relatively shallow water. The water depth where the experiment was conducted is around 360 m. Data were recorded on a 100-m, eight-element vertical array deployed from the Floating Instrument Platform (FLIP) and four bottom-mounted seismometers deployed in an area covering approximately three square kilometers. During the course of the 7-day experiment continuous recording of the ocean environment was made. The recordings contain numerous blue and fin whale calls. Matched field processing was used on the vertical array data to localize and track singing whales. The differences between different animal calls (extended, low frequency calls in the case of blue whales and short, impulsive calls in the case of fin whales) are exploited to track different animal species. Animals were also independently tracked by comparing the predicted (computed using a propagation model) and measured difference in time of arrival recorded in each seismometer pairs. The tracking results obtained from the two techniques are compared.

Session 4aAO

Acoustical Oceanography and Animal Bioacoustics: D. Van Holliday Special Session on Acoustical Measurements of Marine Organisms I

Orest I. Diachok, Chair
3272 Fox Mill Road, Oakton, Virginia 22124

Chair's Introduction—8:30

Invited Papers

8:40

4aAO1. The beginning of Holliday's underwater bioacoustics phase. Paul E. Smith (SIO, Univ. of California, San Diego, CA 92093)

When Van Holliday entered graduate school, research on and management of coastal schooling fishes were in a crisis of disarray. Lasker's sardine physiological work had yielded the unlikely result that the adult sardine had, at one time in the 1930s, consumed most of the production of the California Current. In the analysis of possible errors we had to know: (1) the swimming speed; (2) the size distribution; and (3) the depth distribution of fish in the schools. There were also opportunities. The new DAVID STARR JORDAN, a 50-m research stern trawler, had a Norwegian polished array of WWII tube-type acoustics with steerable powerful sonars at 11 and 30 kHz. This author had equipped the vessel with borrowed trawls. UCSD/SIO/MPL Professors Vic Anderson and Fred Spiess recommended that we add a research program to study the frequency content of acoustic signals. They also nominated a graduate student, Holliday, to design and lead this fisheries research. Van worked at a furious pace to assemble the sources and processors to carry out this work and then doubled and tripled that pace in the 20 months of sea work. Implementation of Holliday's surveys was trapped in the technological dead zone of NOAA.

9:00

4aAO2. Critical scales for understanding the structure, dynamics, and impacts of zooplankton patches. Percy Donaghay (Grad. School of Oceanogr., Univ. of Rhode Island, Narragansett, RI 02882)

The development and application of high resolution acoustic systems by Holliday and Greenlaw has played a key role in demonstrating that zooplankton could can form highly concentrated layers ranging in thickness from a few decimeters to a few meters. These thin layers are remarkable in that they persist longer and have higher spatial coherence than multi-meter thick patches that can be sampled by nets and conventional acoustics. These results are inconsistent with the classical theory that thicker patches are more persistent and thus more important. Simultaneous measurements of temporal and spatial changes in the finescale physical, optical, chemical structure have provided important insights into the mechanisms controlling not only the dynamics of thin zooplankton layers but also their impacts on marine ecosystems and the transmission of sound in the ocean. [Work supported by grants from the ONR Biology and Chemistry program.]

9:20

4aAO3. A tribute to Van Holliday: Completing the circle from plankton to whales. William C. Cummings (5948 Eton Ct., San Diego, CA 92122-3203, oshundbs@san.rr.com)

After the Bachelor's and Master's degrees from the University of Texas and a Doctorate from the University of California, San Diego, Van Holliday hit the water with both feet moving as fast as possible, nary sinking a bit to this day. From the outset he has been a leader and innovator in methods of high-frequency sonar to identify and assess individual animals, populations, and biomasses from zooplankton to adult fishes. Although we'll revisit some of those highlights, my 35-year association with this distinguished scientist also included an opportunity to share comparatively little known acoustics research at the highest marine trophic level, the pinnipeds and whales. Van's scientific endeavors, together with extremely talented colleagues, remain tireless yet somehow he always finds wherewithal to unselfishly help others as testimony to the adage that whenever help is needed one had best seek out their busiest acquaintance. In recognition of his contributions to basic and applied marine science, the U.S. Navy recently bestowed upon Van its highest award for Science and Technology. It will be a pleasure and honor to tell you more of this valued friend and colleague, because you will never hear about it from the scientist, mentor, and gentleman himself.

9:40

4aAO4. Vertical array measurements of humpback whale songs. Whitlow W. L. Au, Marc O. Lammers (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, wau@hawaii.edu), Adam A. Pack, and Louis Herman (Kewalo Basin Marine Mammal Lab., Honolulu, HI 96814)

The songs of eight male humpback whales were recorded at ranges varying from 20 to 40 m with a vertical array of hydrophones that had a flat frequency response to 24 kHz. The songs consisted of bursts of sounds called units. Units were organized into phrases and phrases into themes. Most of the units had mean duration between 1 and 2 s and mean silent periods between units between 1 and

2 s. Many of the recorded songs contained units that had high-frequency harmonics that extended beyond 22 kHz. These harmonic results suggest that humpback whale songs have a broadband quality not previously reported and may provide some insights on the high-frequency limit of hearing in these whales. The source levels of the songs were also estimated by considering the root-mean-square sound-pressure level referenced to 1 m for the unit with the largest level for different phrase within a song. Source levels varied between 171 to 189 dB *re*: 1 μPa. Singing escorts have been regularly observed within two whale lengths of females and these observations and knowledge of source levels provide estimates of sound-pressure levels that male humpback whales expose female whales to.

10:00–10:15 Break

10:15

4aAO5. Van Holliday—A role model for leadership in fisheries acoustics. William A. Karp (NOAA Fisheries, Alaska Fisheries Sci. Ctr., 7600 Sand Point Way NE, Seattle, WA 98115)

Van Holliday completed his Ph.D. dissertation, entitled ''Resonance and Doppler Structure from Pelagic Fish Schools,'' in 1972. He soon published two journal articles based on this research, marking the first steps in a long, distinguished and ongoing contribution to fisheries science. He has published extensively to document his acoustic research on marine life, covering the size spectrum from plankton to marine mammals. This entire body of work is of interest to fisheries biologists who strive to understand the dynamics of populations at all trophic levels to help them better understand the fish populations they study. From this perspective, his advances in acoustic methods, technologies, and instrumentation, and his extensive biological and ecological research should all be considered important contributions to fisheries science. Trained as a physicist, his unique ability to transcend the barriers between physical and biological scientists has been elemental to his success. We recognize Van Holliday for his groundbreaking acoustic research on fish and other forms of marine life, and his scientific and technical excellence. We also recognize Van for his leadership, encouragement, mentoring, and support of colleagues and young scientists, and the vision and focus that he continues to bring to the field of fisheries acoustics worldwide.

10:35

4aAO6. Light and lunar cycle as cues to diel migration of a sound-scattering layer. Kelly J. Benoit-Bird and Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734, benoit@hawaii.edu)

The Hawaiian mesopelagic boundary community is an island-associated midwater scattering layer comprised of small fishes, shrimps, and squids that undergoes diel vertical as well as horizontal migrations. It has been hypothesized that light levels are an important cue or trigger for vertical migration and presumably, horizontal migration. The migration pattern of the scattering layer was measured over complete lunar cycles while the incident light levels were recorded. Due to differences in the rise and set times of the moon and cloud cover, light and lunar cycle were not completely coupled, allowing separation of the light effects of moon phase and other cues associated with lunar cycle. Four calibrated echosounder moorings were deployed with approximately even spacing, perpendicular to the leeward coast of Oahu. Moorings were deployed for one complete lunar cycle at each of three locations, recording 10 echoes every 15 min. Light sensors measured the nocturnal light intensity at 30-s intervals. Statistical analysis revealed significant effects of both light and other lunar cycle cues. Overall, the effect size was very low considering the light transmission characteristics of the subtropical Pacific, making measurement from stationary acoustic platforms critical.

Contributed Papers

10:55

4aAO7. Classification of bioabsorption lines at low frequencies: The Van Holliday connection. Orest Diachok, Stephen Wales (Naval Res. Lab., Washington, DC 20375, OrestDia@aol.com), and Paul Smith (Southwest Fisheries Sci. Ctr., La Jolla, CA 92038-0271)

We will describe preliminary results of the second Bioacoustic Absorption Spectroscopy experiment in the Santa Barbara Channel (SBC), BAS II, which was designed to measure the effects of bioabsorptivity on transmission loss (TL) at frequencies between 0.25 and 6 kHz between an ultrabroadband source and a 16-element vertical receiving array. A fisheries echo sounder and trawls provided bioacoustic parameters. Colocated cores and chirp sonar provided geoacoustic parameters (Turgut, unpublished). Temporal changes in temperature structure were measured with three orthogonal thermistor strings. Highest losses were observed at night when fish were dispersed, and at the average depth of the absorption layer, 13 m, in accord with echo-sounder data and simulations. Absorption lines were evident at the resonance frequencies of 15-cm sardines (1.1 kHz) and 11-cm anchovies (1.8 kHz), in accord with trawls, and 20-cm jack mackerel (0.7 kHz), a species known to inhabit the SBC. This classification is consistent with Holliday's (1972) backscattering measurements of the resonance frequency of jack mackerel (samples collected by Purse Seiner corroborated his classification). Frequencies of absorption lines changed in accord with depth changes at twilight. [Work supported by the ONR and the Southwest Fisheries Science Center.]

11:10

4aAO8. Bioacoustics and D. V. Holliday: A scientific tribute. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

A number of Holliday's seminal contributions to the acoustics of observing aquatic organisms are noted. Schooling pelagic fish have been sized from their echo resonance structure; fish swimming speeds have been measured through the Doppler effect; and snappers/groupers have been enumerated by high-resolution sidescan sonar. Zooplankton have been sized through multiple-frequency measurements of backscatter, and related to the underlying oceanography. Bowhead whales have been localized and tracked by means of a passive sonobuoy array. Certain themes have recurred, emphasizing the importance of both frequency diversity and spatial resolution in quantitative applications. Devices have been designed and built to solve specific problems, advancing the state of the art in these areas. For the student, the respective published work exemplifies the scientific method, introducing an important observational problem, describing its systematic approach, and interpreting results by means of simple but sound physical models. [Work supported by ONR.]

Session 4aBB

Biomedical Ultrasound/Bioresponse to Vibration: Ultrasonic Imaging and Therapy

Pierre Mourad, Chair
Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105

Contributed Papers

8:00

4aBB1. 2D focal-field aberration dependence on time/phase screen position and correlation lengths. Sven Peter Näsholm (Medisinsk teknisk forskningssenter, Dept. of Circulation and Medical Imaging, N-7489 Trondheim, Norway, peter.nasholm@medisin.ntnu.no)

For high-frequency annular array transducers used in medical ultrasound imaging, aberrations due to tissue and body wall have a significant effect on energy transfer from the main lobe to the sidelobes of the acoustic field: that is, the aberrations make the total sidelobe level increase. This effect makes the ultrasound image poor when imaging heterogeneous organs. This study performs an analysis of the focal-field quality as a function of time/phase screen z position and time/phase screen correlation length. It establishes some rules of thumb which indicate when the focal-field sidelobe energy is at its highest. It also introduces a simple screen-scaling model which is useful as long as the screen position is not closer to the focus than a certain limit distance. The scaling model allows the real screen at a depth $z=z_{\text{screen}}$ to be treated as a scaled screen at the position $z=z_{\text{transd}}$. 2D sound fields after 3D propagation from the annular arrays to the focal plane have been simulated using an angular spectrum method. The aberrators are represented by amplitude and phase/time screens.

8:15

4aBB2. Computational evidence for a discrete-scatterer aberration model in medical ultrasound. James C. Lacefield (Univ. of Western Ontario and Robarts Res. Inst., London, ON N6A 5B9, Canada, jlacefield@eng.uwo.ca)

Many techniques for correcting ultrasound focus distortion model the aberrating properties of tissue with a single time-shift screen, but simulations and phantom studies suggest single-screen models are ineffective for transmit focus compensation. Extension of the models to include multiple parallel screens is a logical increment in complexity, but the number of screens must be manageable and readily determined to yield practical aberration correction methods. To assess the feasibility of multi-screen strategies, simulations were performed to search for a general form for the aberration profile of breast tissue. Two-dimensional propagation of 3-MHz planar wavefronts through digitized breast specimens was computed using a *k*-space method [Tabei *et al.*, J. Acoust. Soc. Am. **111**, 53–63 (2002)] and waveforms were sampled at 1-mm intervals along the propagation direction. Arrival time, amplitude, and coherence fluctuations were correlated with scattering from distinct structures. This observation was most apparent when the first derivatives of those parameters with respect to the propagation direction were compared with the connective tissue architecture in the specimens. The assumption underlying time-shift screen models that aberration arises from smooth fluctuations in the acoustic properties of tissue merits reexamination. [Research supported by an NSERC Discovery Grant.]

8:30

4aBB3. Eigenfunction analysis of stochastic backscatter for aberration correction in medical ultrasound imaging. Trond Varslot, Eirik Mo, Harald Krogstad, and Bjørn Angelsen (Norwegian Univ. of Sci. and Technol., Trondheim, Norway)

A filter for aberration correction in medical ultrasound imaging is presented. The filter is optimal in the sense of maximizing the expected energy in a modified beamformer output of the received acoustic backscatter. The situation considered is frequently found in applications when imaging organs through a body wall: aberration is introduced in a layer close to the transducer, and acoustic backscatter from a scattering region behind the body wall is measured at the transducer surface. The scattering region consists of scatterers randomly distributed with very short correlation length compared to the acoustic wave length of the transmit pulse. The scatterer distribution is therefore assumed to be δ-correlated. Theoretical considerations imply that maximizing the expected energy in a modified beamformer output signal naturally leads to eigenfunctions of a Fredholm integral operator, where the associated kernel function is a spatial correlation function of the received stochastic signal. Aberration characterization and aberration correction have been studied for simulated data constructed to mimic aberration introduced by the abdominal wall. The results compare well with what is obtained using a diffraction limited time-reversal filter based on simulated point source data.

8:45

4aBB4. Iteration of ultrasound aberration correction methods. Svein-Erik Maasoey, Bjoern Angelsen (Dept. of Circulation and Imaging, Norwegian Univ. of Sci. and Technol., Olav Kyrres Gate 3, 7491 Trondheim, Norway, svein-erik.masoy@medisin.ntnu.no), and Trond Varslot (Norwegian Univ. of Sci. and Technol., 7491 Trondheim, Norway)

Aberration in ultrasound medical imaging is usually modeled by time-delay and amplitude variations concentrated on the transmitting/receiving array. This filter process is here denoted a TDA filter. The TDA filter is an approximation to the physical aberration process, which occurs over an extended part of the human body wall. Estimation of the TDA filter, and performing correction on transmit and receive, has proven difficult. It has yet to be shown that this method works adequately for severe aberration. Estimation of the TDA filter can be iterated by retransmitting a corrected signal and re-estimate until a convergence criterion is fulfilled (adaptive imaging). Two methods for estimating time-delay and amplitude variations in receive signals from random scatterers have been developed. One method correlates each element signal with a reference signal. The other method use eigenvalue decomposition of the receive cross-spectrum matrix, based upon a receive energy-maximizing criterion. Simulations of iterating aberration correction with a TDA filter have been investigated to study its convergence properties. A weak and strong human-body wall model generated aberration. Both emulated the human abdominal wall. Results after iteration improve aberration correction substantially, and both estimation methods converge, even for the case of strong aberration.

9:00

4aBB5. Enhanced detection of acousto-photonic imaging signals using a photorefractive crystal based system. Lei Sui, Todd Murray, Gopi Maguluri (Dept. of Aerosp. and Mech. Eng., Boston Univ., Boston, MA 02215, suilei@bu.edu), Alex Nieva, Florian Blonigen, Charles DiMarzio (Northeastern Univ., Boston, MA 02115), and Ronald A. Roy (Boston Univ., Boston, MA 02215)

Acousto-photonic imaging is a dual-wave sensing technique where diffuse coherent light wave interacts with a superimposed acoustic field. A phase-modulated light wave emanates from the interaction region and carries with it information about the local opto-mechanical properties of the insonated media. The coherent nature of the light produces a speckle field. The modulation of the speckle field is spatially incoherent, yielding a small modulation depth when collecting multiple speckles or extremely low light levels when detecting at the single speckle level. We report preliminary results from a new detection scheme where the scattered laser light is mixed with a reference beam in a photorefractive crystal. The crystal serves as a dynamic holographic medium and a photorefractive grating is formed from which the reference beam diffracts. The diffracted reference beam has the same spatial structure as the scattered light, and the two interfere at the photodetector where the phase modulation on the scattered beam is converted to an intensity modulation. Measurements of the signals are presented for gel phantoms seeded with suspended polystyrene beads. The results exhibit qualitative agreement with a simple theoretical model. [Work supported by the Center for Subsurface Sensing and Imaging Systems via NSF ERC Award No. EEC-9986821.]

9:15

4aBB6. Combination of *B*-mode imaging and acousto-photonic sensing using a commercial ultrasound scanner. Emmanuel Bossy, Lei Sui, Todd W. Murray, and Ronald A. Roy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ebossy@bu.edu)

The acousto-photonic imaging (API) of a turbid medium is based on the interaction of multiply scattered coherent laser light with an ultrasonic field. The two waves mix and the photons emanating from the interaction region are phase modulated at the ultrasound frequency. This technique yields information on the optical and acoustical properties of the medium, with the interaction region defined by the dimension of the ultrasonic beam. We investigated the feasibility of combining conventional *B*-mode ultrasound imaging and API *in vitro*, using a commercial medical scanner (Analogic AN2300). The AN2300 was used to both generate the *B*-mode images and excite API signals. Gel-based acousto-optic phantoms were fabricated; these contained imbedded targets possessing acoustical and/or optical contrast. Analogous to power Doppler measurements, *B*-mode images were first acquired and then used to select regions of interest within which API signals were generated. API information was then color-coded and superimposed on top of the frozen *B*-mode image. Preliminary results show that API signals can be excited using a commercial scanner, and serve to augment conventional *B*-mode images with information related to the opto-acoustic properties of the medium. [Work supported by the Center for Subsurface Sensing and Imaging Systems via NSF ERC Award Number EEC-9986821.]

9:30

4aBB7. Measurements of the stability of the effective apodization for the nonlinearly generated second harmonic as a function of propagation distance. Russell Fedewa, Kirk Wallace, Mark Holland (Washington Univ. in St. Louis, St. Louis, MO 63130), James Jago, Gary Ng, Matthew Rielly, Brent Robinson (Philips Medical Systems, Bothell, WA), and James Miller (Washington Univ. in St. Louis, St. Louis, MO)

The concept of an effective apodization was introduced to approximate the nonlinearly generated second harmonic field pattern based solely on the linear propagation of a field transmitted at the frequency of the harmonic. We have previously demonstrated that transmitting with an effective apodization determined from measurements made only in the focal plane yields an accurate description of the nonlinearly generated field over a very wide range of depths. The goal of this work was to determine the stability of the effective apodizations obtained from measurements of the ultrasonic field in a series of planes before and beyond the focal zone. Transverse 2D scans of the transmitted fields were performed with a 0.6-mm-diam. hydrophone for both vascular and cardiac arrays. Linear angular spectrum backpropagation of the measured fields determined the effective apodizations. These measurements were compared with simulations based on a nonlinear Burgers equation enhanced angular spectrum approach. The resulting effective apodizations of the second harmonic field were remarkably constant (differing by typically a few percent and never by more than 9.4% from the mean) for ten axial positions ranging from 0.1 to 1.6 times the focal distance. [Work supported by NIH HL072761 and Philips Medical Systems.]

9:45

4aBB8. Dual transmission model of the fetal heart tone. Donald A. Baker (Ste. 230, 3003 3rd Ave. NE, Camas, WA 98607) and Allan J. Zuckerwar (NASA Langley Res. Ctr., Hampton, VA 23681)

Detection of the fetal heart tone by auscultation is sometimes easy, other times very difficult. In the model proposed here, the level of difficulty depends upon the position of the fetus within the maternal abdomen. If the fetus lies in the classical left/right occiput anterior position (head down, back against the maternal abdominal wall), detection by a sensor or stethoscope on the maternal abdominal surface is easy. In this mode, named here the "direct contact" mode, the heartbeat pushes the fetus against the detecting sensor. The motion generates pressure by impact and does not involve acoustic propagation at all. If the fetus lies in a persistent occiput posterior position (spine-to-spine, fetus facing forward), detection is difficult. In this, the "fluid propagation" mode, sound generated by the fetal heart and propagating across the amniotic fluid produces extremely weak signals at the maternal surface, typically 30 dB lower than those of the direct contact mode. This reduction in tone level can be compensated by judicious selection of detection frequency band and by exploiting the difference between the background noise levels of the two modes. Experimental clinical results, demonstrating the tones associated with the two respective modes, will be presented.

10:00–10:15 Break

10:15

4aBB9. Flow measurements based on speckle decorrelation: Simulation and experiment. Oliver D. Kripfgans, Juan Zhu, Jonathan M. Rubin, J. Brian Fowlkes (Dept. of Radiol., Univ. of Michigan Health Systems, Ann Arbor, MI 48109-0553), and Anne L. Hall (GE Med. Systems, Milwaukee, WI)

Traditional Doppler-based flow measurements suffer from bad signal to noise for large angles between the wavevector and the flow direction. To overcome this limitation, speckle decorrelation might be used to measure lateral flow. Experiments were performed on a flow phantom (tube diameter 6.35 mm, flow 1.6 mL/min) with the tube axis positioned in the imaging plane of a GE Logiq 9 scanner. IQ data sets of ten frames with 16 firings per scanline were recorded and speckle decorrelation used to estimate flow speeds throughout the image. The decorrelation computations were performed over different kernel types and compared to simulations performed using Field II by J. Jensen with acoustical transmit as well as beamforming parameters set to match experiments. Speckle decorrelation rates in experiments scale correctly for the parabolic flow profile inside the tube. Simulations reproduced a similar but smoother profile. Flow velocities could be estimated using a scaling factor based on the spatial correlation of the beam. The combination of velocity estimates from Doppler and speckle decorrelation may provide a more uniform display of flow and lead to less angle dependence. [Research supported by U.S. Army Grant No. DAMD17-00-1-0344 and GE Medical Systems.]

10:30

4aBB10. Validation of Doppler ultrasound measurements using particle, image velocimetry in a flow phantom. John Cosgrove (School of Phys., Univ. of Edinburgh, Edinburgh EH9 3JZ, UK), Siobhan Meagher, Peter Hoskins, Clive Greated (The Univ. of Edinburgh, UK), and Richard Black (The Univ. of Liverpool, UK)

Cardiovascular disease is responsible for over 50% of all deaths in the world and there is a substantial amount of evidence which suggests that abnormal vessel wall shear stress is correlated with the development of atherosclerosis. Wall shear stress is calculated from wall shear rates, the measurement of which is a technically challenging problem for ultrasound. In this study a flow phantom consisting of a meshed-gear pump and corresponding control electronics is used to generate a range of flow waveforms in a straight tube. These flows are measured using Doppler ultrasound and compared to corresponding particle image velocimetry (PIV) measurements and to analytical solutions of the flow equations for a range of Wormersley parameters. Although previous studies have been undertaken calibrating Doppler ultrasound in straight tubes, they have not used PIV. This study serves as a prelude to investigations using PIV to assess the accuracy of Doppler ultrasound in phantoms with anatomically realistic geometries for which there are no analytical solutions to the flow. [Research funded by the Engineering and Physical Sciences Research Council UK.]

10:45

4aBB11. *In vitro* microscopic imaging of rt-PA thrombolysis with 120-KHz ultrasound in a human clot model. Jason Y. Cheng (Dept. of Emergency Medicine, Univ. of Cincinnati, 231 Albert Sabin Way, Cincinnati, OH 45267-0769, Jason.Cheng@uc.edu), Christy K. Holland (Univ. of Cincinnati, Cincinnati, OH 45267-0586), and George J. Shaw (Univ. of Cincinnati, Cincinnati, OH 45267-0769)

Substantial enhancement of recombinant tissue plasminogen activator (rt-PA) thrombolysis can be achieved with diagnostic frequency ultrasound. Microscopic visualization of human whole-blood clots treated with human fresh frozen plasma (HFFP), rt-PA, and 120-KHz pulsed ultrasound offers a useful method to study the possible mechanisms. Whole human-blood clots were formed inside of glass micropipette (1.7-mm diameter) and placed in a water tank at 37 °C. The clot–plasma interface was imaged using an inverted optical microscope. Clots were exposed to HFFP (control), HFFP and rt-PA (0.0945 mg/ml), (sham), or HFFP, rt-PA (0.0945 mg/ml), and 120-KHz pulsed ultrasound (ultrasound treated) for 30 min. Thrombolysis at the clot surface was recorded with a CCD camera and the lytic front was analyzed as a function of time. Images of the clots were analyzed to quantify the overall amount of lysis and compared to a dynamic chemical model of enzymatic degradation in the presence of ultrasound-induced flow. Sham and ultrasound-treated clots showed significant lysis with the lytic rate of (1.33 ± 0.02 microns/min; mean $\pm$ s.d.) and (5.17 ± 0.10 microns/min), respectively. Ultrasound enhanced thrombolysis by a factor of 4. In addition, the lytic rate profile may offer a suggestion to the mechanism of enhancement. No thrombolysis was observed in control clots. [The authors of this study gratefully acknowledge the support of the Whitaker Foundation Grant RG-0128-01.]

11:00

4aBB12. Ultrasonic brain therapy: First trans-skull *in vivo* experiments on sheep using adaptive focusing. Mathieu Pernot, Jean-Francois Aubry, Michael Tanter, Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.M.R. C.N.R.S. 7587, 10 rue Vauquelin, 75005 Paris, France, jf.aubry@espci.fr), Anne-Laure Boch, and Michèle Kujas (Hôpital de la Salpêtrière, 75013 Paris, France)

A high-power prototype dedicated to trans-skull therapy has been tested *in vivo* on 20 sheep. The array is made of 200 high-power transducers working at 1-MHz central and is able to reach 260 bars at focus in water. An echographic array connected to a Philips HDI 1000 system has been inserted in the therapeutic array in order to perform real-time monitoring of the treatment. A complete craniotomy has been performed on half of the treated animal models in order to get a reference model. On the other animals, a minimally invasive surgery has been performed thanks to a time-reversal experiment: a hydrophone was inserted at the target inside the brain thanks to a 1-mm^2 craniotomy. A time-reversal experiment was then conducted through the skull bone with the therapeutic array to treat the targeted point. For all the animals a specified region around the target was treated thanks to electronic beam steering. Animals were finally divided into three groups and sacrificed, respectively, 0, 1, and 2 weeks after treatment. Finally, histological examination confirmed tissue damage. These *in vivo* experiments highlight the strong potential of high-power time-reversal technology.

11:15

4aBB13. High-resolution ultrasonic brain imaging: Noninvasive adaptive focusing based on twin arrays. Francois Vignon, Jean-Francois Aubry, Michael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.M.R. C.N.R.S. 7587, 10 rue Vauquelin, 75005 Paris, France, jf.aubry@espci.fr)

Ultrasonic imaging systems assume a constant acoustic velocity in human tissues in the beamforming process. However, in the case of brain imaging, strong skull aberrations induce a displacement of the focal stain, a spreading in the main lobe, and an increase in the sidelobes level. This limits considerably ultrasonic brain imaging applications. It has been shown that a very accurate focusing could be achieved through a human skull by using an inverse filter technique. This method, based on a set of acoustic sensors located inside the brain, demonstrated invasively that it was possible to achieve high-resolution brain imaging. A noninvasive ultrasonic method is presented. It is based on two arrays located on each side of the head. Each array focuses on the other one to deduce the aberrations induced by the skull bone. Once the effect of the skull bone in front of each array is extracted, the wavefronts are corrected and emitted. This noninvasive technique restores the position of the focal stain and lowers the secondary lobe level up to 15 dB compared to a cylindrical focusing.

11:30

4aBB14. Ultrasound produced homogenization and liquefaction of brain tissue *in vivo* at pressures beyond the cavitation threshold. Natalia Vykhodtseva, Nathan McDannold, and Kullervo Hynynen (Havard Med. School/Brigham and Women's Hospital, 221 Longwood Ave., Rm. 013, Boston, MA 02115)

Histological effects of short-pulsed focused ultrasound at pressure amplitudes above the cavitation threshold were studied in the brain. Sixty locations in the brains of 13 rabbits were sonicated at a depth of 4–10 mm (1.5 MHz, burst lengths, 1–100 ms, PRF 1–100 Hz, acoustic power, up to 950 W). MRI was used to aim the beam, monitor the temperature, and evaluate the effects. It was found that high-intensity short-pulsed ultrasound could effectively destroy targeted tissues. The temperature elevation measured during the sonication was below the threshold for tissue damage. The tissue effects were different from that which has been described for the thermal or cavitation mechanisms. The lesions appeared as homogeneous regions with no distinguishable cell debris. The central core of the lesion region consisted of a pudding-like mass, arranged in a helically coiled, layered fashion. In some cases, the core seemed to be completely liquefied. These findings appeared to be associated with the nonlinear ultrasound absorption and possibly cavitation-related effects. The spiral shape observed in some cases may indicate the existence of an axial radiation torque in the focal zone. The lesions' appearance may point to new interactions between ultrasound and tissue, which could be exploited for new therapies.

11:45

4aBB15. Transcranial MRI-guided FUS-induced BBB opening in the rat brain. Lisa H. Treat, Nathan J. McDannold, and Kullervo Hynynen (Dept. of Radiol., Brigham and Women's Hospital/Harvard Med. School, 221 Longwood Ave., LMRC Rm. 013, Boston, MA 02115, treat@mit.edu)

The blood-brain barrier (BBB) has been a major limitation in treating diseases of the brain because therapeutic agents are either unable to penetrate or have dose-limiting side effects in diffuse opening of the BBB. A previous study demonstrated that focused ultrasound (FUS) can locally

open the BBB in a rabbit model when a piece of skull is removed and that magnetic resonance imaging (MRI) can be used to guide and monitor the procedure. This study examined whether the same desired effect of local BBB disruption can be achieved by applying FUS through an intact skull in a rat model. Twenty-eight Sprague-Dawley rats were anesthetized, shaved, and sonicated at four focal locations in the brain, using a 1.5-MHz focused transducer. Contrast-enhanced MR images were obtained before and after sonication. The images indicated contrast agent penetration at the focal coordinates following Optison-enhanced sonication. This study demonstrated that the distortion of the ultrasound beam by the rat skull was not significant enough to inhibit focal BBB opening. Subsequent experiments using MRI-guided FUS to aid in targeted drug delivery to brain tumors in a rodent model could thus be performed more efficiently without cranial surgery. [Research funded by NIH Grant No. CA76550.]

Session 4aEA

Engineering Acoustics: Transducers, Arrays and Calibration

Kirk Jenne, Cochair
NAVSEA Division Newport, 1176 Howell Street, Newport, Rhode Island 02841-1708

Kenneth Walsh, Cochair
K M Engineering Ltd., 51 Bayberry Lane, Middletown, Rhode Island 02842

Chair's Introduction—8:25

Contributed Papers

8:30

4aEA1. A new microsystem for near whispering. Sungjoon Choi, Wonkyu Moon (Dept. of Mech. Eng., POSTECH, San 31, Hyoja-Dong, Nam-gu, Pohang, Kyungbuk 790-784, South Korea), and Jeong Hyun Lee (iCurie Lab., Daechi-dong, Kangnam-gu, Seoul 135-846, South Korea)

A new microphone system was developed to monitor the human voice near the microphone in a noisy environment. The system is equipped with two special functions in addition to the usual microphone functions: reduction of air blow effects by the mouth and focused reception to a sound source. A wind filter was developed to reduce the air blow effects from the mouth during speaking. This filter is a plate perforated by an array of small holes; the method used to design the filter is also presented. To achieve focused reception, four microphones were used in conjunction with a new signal processing method. The proposed signal processing method effectively increases the directivity in the desired direction. Additionally, it provides the system with focusing on the source since the source is located adjacent to the system. A prototype of the proposed system was fabricated and subjected to performance tests. The results showed that air blow effects can be reduced by up to 20 dB and the directional gain is more than 4 dB. The proposed microphone system shows such good performance that it can be used in mobile phones for whispering communication.

8:45

4aEA2. On the potential performance of micro-fabricated subminiature microphones based on electron surface tunneling. Michael Pedersen (CNRI, 1895 Preston White Dr., Ste. 100, Reston, VA 20191)

The principle of electron tunneling across a vacuum potential energy barrier, as utilized in surface tunneling microscopy (STM), is a well known, highly sensitive, method for the detection of surface features and displacements in the subnanometer regime. In this paper, the basic properties of the surface tunneling principle and its possible utilization in acoustic sensors are discussed, with particular focus on the expected noise properties of a microphone based on tunneling in comparison to more conventional devices based on capacitive detection. It is found that the noise performance is largely limited by thermal noise due to the acoustical radiation impedance of the device in combination with electronic noise in the circuit used to establish and control the tunneling current. Finally, possible methods for sensor fabrication are discussed with special attention aimed at resolving the common problem of vibration sensitivity associated with STM. This problem stems from the relative large inertial mass and low flexural rigidity of the tunneling tip suspension system in most detector designs. Alternative approaches suitable for acoustic sensors in which the tunneling tip is fixed are described and it is shown that low vibration sensitivities comparable to capacitive devices can be achieved.

9:00

4aEA3. Numerical evaluation of the omnidirectional behavior of regular polyhedron loudspeakers. Heather Smith and Timothy Leishman (Dept. of Phys., Brigham Young Univ., N283 ESC, Provo, UT 84602, hm73@email.byu.edu)

A regular polyhedron loudspeaker (RPL) consists of a rigid enclosure in the shape of a regular polyhedron with loudspeaker drivers mounted within each face. Room acoustics measurements often incorporate dodecahedron loudspeakers presumably as omnidirectional sources of sound. This research is intended to support experimental work to determine which of the five regular polyhedron loudspeaker configurations actually produces the most omnidirectional field over a wide frequency range. Loudspeaker models based on the boundary element method were created to assess frequency-dependent radiation patterns. These models assume that the loudspeaker drivers radiate as plane circular pistons. Far-field sound pressures were generated for several field points around the models (5-deg polar and azimuthal angle increments). Results are presented as directivity balloon plots and compared with experimental measurements. Other numerical results are shown as variations on the experimental setup. These involve dimensional scaling of an entire loudspeaker and dimensional scaling of driver diameters while keeping enclosure dimensions constant. They also involve generation of data at more closely spaced field points (1-deg polar and azimuthal angle increments). Results from a spherical enclosure with radiating pistons centered on the angular coordinates of dodecahedron facial centers are also compared to those produced by a dodecahedron loudspeaker model.

9:15

4aEA4. Optical mapping of the acoustic output of a focused transducer. Robert D. Huber, Diane J. Chinn, and David H. Chambers (Lawrence Livermore Natl. Lab., P.O. Box 808 L-333, Livermore, CA 94551)

A Michelson interferometer is used to map the ultrasonic displacement of the lens at the end of a delay rod of a 50-MHz immersion transducer. The purpose of mapping the displacement is to provide a source function to a model that predicts the ultrasonic propagation in, and interaction with, various materials. The output of the Michelson interferometer can be calibrated, and then used to determine the displacement of the transducer lens surface moving at ultrasonic frequencies. Using the interferometer, the displacement of the transducer lens is measured at discrete points along its surface. This displacement map then provides the ultrasound propagation model with the actual source function. Direct comparison between a model with a simulated source function and experimentally obtained data is presented. [Work performed under auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under Contract No. W-7405-ENG-48.]

9:30

4aEA5. Effects of coupled vibrations in cylindrical shell transducers. Boris Aronov, Sundar Regmi, and David A. Brown (BTech Acoust. and Univ. of Massachusetts Dartmouth, Acoust. Res. Lab., ATMC, 151 Martine St., Fall River, MA 02723, DBrown122@cox.net)

Coupled vibrations in cylindrical transducers can have unwanted effects on the resulting radiation or reception of sound. An investigation of vibrations of piezoelectric cylindrical shell transducers using analytical and experimental methods shows the presence of an additional flexural resonant mode that was not previously reported [see J. F. Haskins and J. L. Walsh, ''Vibrations of ferroelectric cylindrical shells with transverse isotropy,'' J. Acoust. Soc. Am. **29**, 6 (1957)]. The additional flexural mode can be strongly coupled to radial and axial modes in cylindrical shells with finite thickness. The effect can be substantial, resulting in unwanted resonances that may be in the frequency band of interest for both projectors and receivers. An equivalent circuit is derived, which describes the cylindrical transducer operation taking into account the effect of coupled vibrations. The results of calculations in terms of resonant frequencies and frequency responses for both transmit and receive modes are in good agreement with experiments. The overall performance of the cylindrical shell transducer can be greatly enhanced by using cylinders of appropriate height to diameter aspect ratio. [Work supported in part by ONR 321SS.]

9:45

4aEA6. Development of local volume displacement sensors for vibrating plates. Ron Wendt and Marcellin Zahui (Mech. Eng., Univ. of North Dakota, Grand Forks, ND 58202, Marcellin.Zahui@und.nodak.edu)

Development of local volume displacement sensors is presented. This development supports the implementation of noise control techniques that are based on minimization of local volume displacements, velocities, or accelerations of a vibrating structure. In this paper, a general methodology for the development of local volume displacement sensors for vibrating plates using PolyVyniliDene Fluoride (PVDF) is presented. This methodology was verified experimentally for a clamped plate. The local volume displacement measured using a PVDF sensor matched the local volume displacement found using multiple accelerometer measurements. The resulting sensor spans the entire length of the plate and covers the width of the area of interest. The sensor is composed of a number of strips whose width varies quadratically on the local area and linearly elsewhere. Design issues for a clamped plate, a simply supported plate, and a plate with arbitrary boundary conditions are discussed along with a presentation of some sample sensor shapes.

10:00–10:15 Break

10:15

4aEA7. Volume displacement sensors for vibrating beams: Numerical approach. Sai Prassad Jangiti and Marcellin Zahui (Mech. Eng., Univ. of North Dakota, Grand Forks, ND 58202, Marcellin.Zahui@und.edu)

The development of volume displacement sensors for vibrating beams is revisited with emphasis on numerical approach. This development supports the implementation of noise control techniques that are based on minimization of volume displacements, velocities, or accelerations of a vibrating structure. This paper first reviews some of the existing general methodologies for the development of volume displacement sensors for vibrating beams using PolyVinyliDene Fluoride (PVDF). The presentation includes the quadratic and modal development of volume displacement sensors for vibrating beams. These techniques are extended to numerical analysis by discretizing the beam and assuming constant sensor shape on each beam element. The result is a system of linear equations in which the assumed constant shapes are the unknowns. The size of the system of equations, which determines the accuracy of the sensor, is directly related to the highest frequency in the signal to be processed. The resulting sensors are numerically and experimentally verified for a simply supported beam. The results show for low-frequency application, relatively simpler sensor shapes can be utilized, whereas for higher frequencies the sensor shape converges to quadratic function. Finally, sensor accuracy and the range of the application frequency are discussed and some sample shapes presented.

10:30

4aEA8. Effects of mutual impedance on the radiation characteristics of transducer arrays. Haksue Lee, Jinhee Tak, Wonkyu Moon, and Geunbae Lim (Dept. of Mech. Eng., POSTECH, San 31, Hyojadong Namgu, Pohang, Kyungbuk 790-784, South Korea)

The mutual resistance of transducer arrays is investigated in order to design arrays with improved performance for high intensity sounds at a given frequency. This work proposes the theory that the mutual resistance is related to the loading effects of pressure waves propagated from a piston driver on the surface of another driver. Using this interpretation, the important characteristics of the mutual resistance of two piston drivers are explained and the conditions for local maxima in the mutual resistance are easily determined. On the basis of analyses of the interactions between a driver and acoustic pressure waves, we propose a method to determine the driver radius and the distance between two drivers that give maximum mutual radiation resistance. To evaluate the proposed method, the total resistance of a transducer array is calculated using the formulas for mutual and self-resistance established by Pritchard. The results of the calculations of the total resistances of arrays with many drivers show that a transducer array with drivers arranged sparsely can achieve a larger value of the radiation power per unit area as well as better radiation efficiency than an array in which the drivers are in a closely packed arrangement at a given frequency.

10:45

4aEA9. Measurements of mutual radiation impedance in arrays of piezoelectric cylindrical shell transducers. Tetsuro Oishi, Boris S. Aronov, and David A. Brown (Acoust. Res. Lab., Dept. of Elec. and Computer Eng., Univ. of Massachusetts at Dartmouth, MA 02723)

We present our recent experimental results of the mutual radiation impedance of piezoelectric cylindrical shell transducers in arrays using the Z- and V-methods previously reported by the authors [J. Acoust. Soc. Am. **112**, 2407 (2002)]. The mutual radiation impedance in an array of two coaxially aligned cylindrical transducers was determined as a function of separation distance. The results obtained by both methods are in good agreement with the calculations based on the analytical equations for a similar array developed by D. H. Robey [J. Acoust. Soc. Am. **27**, 706–710 (1955)]. The mutual radiation impedances in other array geometries in-

cluding planar and curved arrays of cylindrical shell transducers, in which there is no known analytical solution, were also measured by the same methods in our underwater test facility. [Work supported in part by ONR 321 SS.]

11:00

4aEA10. Audio demodulation of radiation pressure from modulated ultrasound. Minoru Toda (Measurement Specialties, Inc., 460 E. Swedesford Rd., Ste. 3005, Wayne, PA 19087)

Parametric arrays produce audible sound from the audio frequency modulation of high-intensity ultrasound as a result of nonlinear propagation. This paper presents observations of the audible sound resulting from modulated radiation pressure. Audio signals due to the parametric array effect depend strongly on the square of the modulation frequency. It is often shown that the low-frequency region shows a flat frequency response, although it has not been clarified whether this is related to the system noise floor, modulated radiation pressure, or microphone nonlinearity. In this investigation, microphone nonlinearity was found to be very strong at short distances unless acoustic filtering was used. A novel low-pass acoustic filter (passes less than 10 kHz) was constructed using four layers of polyethylene spaced at 4 mm. Using this scheme, the ultrasonic carrier was attenuated and microphone nonlinearity was reduced by six orders of magnitude. Modulated radiation pressure was successfully observed at distances of 1 to 3 m. The observed audio signal showed a flat frequency response and agreed with theoretical predictions for modulated radiation pressure. The reproduced audio signal from radiation pressure appears in a lower frequency region of parametric array audio systems, and the effect should be taken into the system design.

11:15

4aEA11. Infrasonic pistonphone. Allan J. Zuckerwar, John E. Teter, and William E. Robbins (NASA Langley Res. Ctr., M.S. 493, Hampton, VA 23681, a.j.zuckerwar@larc.nasa.gov)

A hand-held, battery-powered pistonphone was designed to calibrate infrasonic microphones in the field. With a battery voltage of 12 V, a prototype pistonphone provides a sound pressure level of 110 ± 0.5 dB *re* 20 μPa at a frequency of 13.8 Hz. The microphone is inserted into an acoustic coupler, in which a stretched diaphragm is excited to generate infrasound. The diaphragm motion is generated by an eccentric ball bearing, such that the outer race does not rotate but remains in contact with the diaphragm at the same point during a complete rotation cycle. Thus this mechanism avoids sliding friction and insures a positive diaphragm displacement, independent of acoustic load. The diaphragm tension is adjusted to a sufficiently high level to prevent "floating" (separation from the eccentric) at the operating frequency. Further, the pistonphone is designed to have a high "thermal wave number" to insure adiabatic as opposed to isothermal wave propagation.

11:30

4aEA12. Method and system for calibrating acoustic receivers in borehole logging tools. Fernando Garcia Osuna and Toru Ikegami (Schlumberger, 110 Schlumberger Dr., Sugar Land, TX 77478, fgarcia@sugar-land.oilfield.slb.com)

Monopole and quadrupole contamination on dipole measurements is one of the main problems for sonic logging tools that employ arrays of acoustic receivers to measure a borehole's dipole mode. The amplitude and phase mismatch of the array of sensors in acoustic logging tools plays an important role in the quality of dipole measurements. Acoustic receivers often have different sensitivities, and different sensitivities to the same wave field result in a greater possibility of no dipole contamination. Even similarly or identically manufactured receivers tend to report different amplitudes and time measurements (i.e., amplitude and phase mismatch). In practice, the no-dipole modes are removed by making the measurements at different azimuth and extracting the dipole mode from those that reject the other contaminating modes. Therefore, to improve rock formation slowness estimation and downhole modal computation, it is necessary to calibrate acoustic logging tools by detecting and correcting the amplitude and phase mismatch of the individual sensors mounted in the logging tools. A method and system for calibrating acoustic receiver arrays mounted in a downhole logging tool and the factors that affect their sensitivity are discussed in this paper.

THURSDAY MORNING, 27 MAY 2004 VERSAILLES BALLROOM, 10:00 A.M. TO 12:00 NOON

Session 4aED

Education in Acoustics: Hands on Demonstrations for High School Students

Uwe J. Hansen, Chair

Physics Department, Indiana State University, Terre Haute, Indiana 47809

Chair's Introduction—10:00

Approximately 20 acoustics demonstrations will be distributed in the room. All demonstrations will be available for high school students' hands-on experimentation. Participation by other conference attenders is welcome as long as their activity does not interfere with student learning.

Session 4aMU

Musical Acoustics: General Topics in Musical Acoustics

Roger J. Hanson, Chair
Physics Department, University of Northern Iowa, Cedar Falls, Iowa 50614

Contributed Papers

8:30

4aMU1. Radiation control applied to sound synthesis: An attempt for "spatial additive synthesis." Olivier Warusfel, Nicolas Misdariis, Terence Caulkins, and Etienne Corteel (IRCAM, 1 place Stravinsky, 75004 Paris, France)

Sound synthesis is generally focused on the reproduction of the spectral characteristics of the source or on the simulation of its physical behavior. Less attention is paid to the sound playback step which generally results in a simple diffusion on a conventional loudspeaker setup. Stating the perceptual importance of a faithful reproduction of the source radiation properties, the paper presents a method combining a synthesis engine, based on physical modeling, with a rendering system allowing an accurate control on the produced sound-field. Two sound-field synthesis models are considered. In the first one, a local 3D array of transducers is controlled by signal processing for creating elementary directivity patterns that can be further combined in order to shape a more complex radiation. The dual approach consists in surrounding the audience with transducer arrays driven by wave field synthesis in order to simulate the sound field associated to these elementary directivity patterns. In both cases, the different radiating modes of a given instrument are synthesized separately, in conjunction with their associated radiation pattern, and then superimposed in the spatial domain, i.e., during the propagation in air. This approach, referred as ''spatial additive synthesis,'' is illustrated, taking the example of different musical instruments.

8:45

4aMU2. Digital filters for accurate simulation of wave propagation losses in tubes. Reiner Wilhelms-Tricarico and Richard McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420, reiner@speech.mit.edu)

A boundary layer approximation for viscous damping in one-dimensional wave transmission in a tube results in an irrational frequency-dependent damping filter for sound propagation. This filter can be approximated with high accuracy by a rational filter function that can be obtained from Padé approximations or continued fraction expansion. Taking into account the viscous losses in a Kelly–Lochbaum structure that represents sound propagation in a tube with spatially varying cross section results in replacing the delay elements of the lattice filter for the loss-free case by special recursive filters. The design, implementation, and applications of the filter structures will be presented.

9:00

4aMU3. In-room sound reproduction using active control: Simulations in the frequency domain and comparison with wave field synthesis. Philippe-Aubert Gauthier, Alain Berry (GAUS, Mech. Eng. Dept., Université Sherbrooke, 2500 boul. Université, Sherbrooke, QC J1K 2R1, Canada, philippe_aubert_gauthier@hotmail.com), and Wieslaw Woszczyk (McGill Univ., Montréal, QC H3A 1E3, Canada)

Active sound control simulations were performed for progressive sound field reproduction over a ''large'' area using multiple monopole loudspeakers. The model is limited to the simulation of the acoustical output of the prescribed loudspeaker array in a simple room, and is based on achieving an optimal control in the frequency domain. This rather simple approach is chosen for this first feasibility study concerning a limited number of possible configurations of sensing microphones and loudspeakers. Other issues of interest concern the comparison with wave field synthesis, the control mechanisms and transducer configurations. As it is demonstrated, in-room reproduction of sound field using active control can be achieved with a residual normalized squared error below 2% while open-loop wave field synthesis gives more than 100% of error in the same situation. Usage of active control technique suggests the possibility to automatically overcome the room's natural dynamics. A special surrounding configuration of sensors is introduced for a sensor-free listening area. [Work supported by NSERC, NATEQ, VRQ, and Université de Sherbrooke.]

9:15

4aMU4. Philosophical and cultural perspectives on acoustics in Vedic Hinduism. M. G. Prasad (Dept. of Mech. Eng., Stevens Inst. of Technol., Hoboken, NJ 07030, mprasad@stevens.edu)

Acoustics plays a very important multi-faceted role in Vedic Hinduism. Vedas, that is an infinitely large collection of chants (mantras) in ancient Sanskrit language, form the foundational literature of Vedic Hinduism. The Vedic chants have specific acoustical qualities and intonations. The Vedic literature describes the various aspects of acoustics, namely, philosophical, spiritual, and cultural. The use of sounds from conch-shell, bells, cymbal in addition to the Vedic chants in rituals shows the spiritual aspects. Vedic literature discusses the role of sound in the philosophical understanding of our world. Music, both vocal and instrumental, plays an important role in the cultural aspects of Vedic Hinduism. It can be seen that certain musical instruments such as ''mridangam,'' a percussion drum, reflect scientific principles underlying in their design. This paper presents an overview of the various important and interesting roles of acoustics in Vedic Hinduism.

9:30

4aMU5. Measurement of effects on tone with lip-protecting music splints for wind instrument players. Chigusa Katada, Kazunori Nozaki, Miharu Imai (Div. of Community Dentistry and Informatics, Molecular Oral Biol. and Dentistry, Osaka Univ., 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan, c_katada@dent.osaka-u.ac.jp), Masayuki Kawamoto, Yuko Shima, Hiroo Tamagawa, Yoshinobu Maeda (Osaka Univ. Dental Hospital, Suita, Osaka 565-0871, Japan), Naoki Ohboshi (Kyoto Univ. Medical Hospital, Sakyo-ku, Kyoto 606-8501, Japan), and Tadao Toda (Osaka College of Music, Osaka 561-8555, Japan)

To protect against lip trauma from wind instruments, music splints that cover the sharp edges of incisor teeth are often manufactured by dentists. Wind instrument players who have installed these custom-made music splints often express not only their lip comfort but also changes in their tone quality. In this study, we investigated the effect of the splints to the tone quality. We recorded three types of trumpet sounds such as long tones, arpeggios with perfect fifth, and tonguing tones with and without using a splint, respectively, by a professional trumpet player in an anechoic room. After fast Fourier transform, the higher harmonics was observed more in the splint group than in the nonsplint group, with sharp peaks from 5000 to 8000 Hz. We also examined the differences of these sound groups with recognition tests by two groups of listeners such as

professional musicians and nonprofessional persons. Though sound-pressure levels of higher harmonics in two sound groups were lower than those at 400 to 2000 Hz, the musically trained persons recognized the difference perfectly. These results suggest the target of measurement to evaluate the effect of music splints.

9:45

4aMU6. The Sagrada Familia Cathedral where Gaudi envisaged his bell music. Shigeru Yoshikawa and Takafumi Narita (Dept. of Acoust. Design, Grad. School of Kyushu Univ., 4-9-1 Shiobaru, Minami-ku, Fukuoka 815-8540, Japan, shig@design.kyushu-u.ac.jp)

The Sagrada Familia Cathedral in Barcelona, Spain was constructed in 1882. According to Antoni Gaudi, who worked over its grand plan, the Cathedral was supposed to be a huge musical instrument as a whole in the event of completion. As as result, the music of bells was expected to echo through the air of Barcelona from the belfries. However, Gaudi's true intention cannot be exactly known because the materials prepared by him were destroyed by war fire. If his idea of the Sagrada Familia as an architechtural music instrument is true, an acoustical balance should be considered between the roles of the Cathedral: bell music from the belfries and quiet service in the chapel. Basic structure of the Sagrada Familia seems to be an ensemble of twin towers. Following such speculation, we made a simplified acrylic 1/25-scale model of the lower structure of a twin tower located at the left side of the Birth Gate. The higher structure of this twin tower corresponds to the pinnacle where the bells should be arranged. The lower structure (about 43 m in actual height) has five passages connecting two towers. One of two towers includes five or six tandem columns whose ends are both squeezed to about 1.5 m in diameter. These columns seem to function as a kind of muffler. The location and shape of the roof over the nave is indefinite and tentatively supposed at the top of the lower structure. Based on our scale model, acoustical characteristics of the lower twin-tower structure as a muffler and acoustical differences between the exterior field and nave field will be reported and discussed.

10:00

4aMU7. Synthesis of audio spectra using a novel Bessel expansion. V. Vijayakumar and C. Eswaran (Multimedia Univ., FOSEE, Melaka 75450, Malaysia)

It is shown in this paper that the intensity variations of an audio signal in the frequency domain can be expressed by using a novel mathematical function containing a series of weighted complex Bessel functions. By proper choice of values for two parameters, this function can transform an input spectrum of discrete frequencies of unit intensity into the known spectra of different musical instruments. Two specific examples of musical instruments are considered for evaluating the performance of this method. It is found that this function yields musical spectra with a good degree of accuracy. The proposed method is compared with known synthesis techniques such as FM, AFM, and DFM [J. M. Chowning, J. Audio Eng. Soc. **21**, 526–534 (1973); Palamin *et al.*, *ibid*. **36**, 671 (1988); Tan, *ibid.* **42** (11), 918–926 (1994)]. A brief discussion on the physical basis for the derivation of the proposed function is also presented.

THURSDAY MORNING, 27 MAY 2004 ROYAL BALLROOM A, 8:35 A.M. TO 12:00 NOON

Session 4aNS

Noise and jointly sponsored by all ASA Technical Committees: On the Occasion of His 90th Birthday: Special Session to Honor the Contributions of Leo L. Beranek to Acoustics and Teaching

Tony F. W. Embleton, Chair

80 Sheardown Drive, Box 786, Nobleton, Ontario L0G 1NO, Canada

Chair's Introduction—8:35

Invited Papers

8:40

4aNS1. Leo Beranek's contributions to architectural acoustics from development of criteria for design to practical applications in buildings of all types. William J. Cavanaugh (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776, wcavanaugh@cavtocci.com)

This paper traces Leo Beranek's continuing contributions in architectural acoustics since the author joined the consulting staff of Bolt Beranek and Newman Inc. in February 1954. They are legion, including room acoustics, sound and vibration transmission, control of ambient noise from building mechanical electrical and plumbing systems, electrocoustic applications and other issues through the evaluation of the finished building spaces themselves. While Beranek's classic texts, *Acoustical Measurements* (1948) and *Acoustics* (1954) served as essential references for the early BBN staff, criteria in architectural acoustics were severely limited if not nonexistent. If he did not take up the challenge to fill a criterion gap, his inspirational leadership guided others to do so, as it did the author and colleagues trying to better understand speech privacy in buildings. Beranek's real labor of love in architectural acoustics is in understanding music performance halls. Beranek's dedicated research over a span of five decades has produced countless journal papers in concert hall acoustics as well as three important texts, the most recent in 2003. Leo Beranek's contributions have earned him the Sabine and Gold Medals of the Acoustical Society and countless other honors including most recently the National Medal of Science.

9:00

4aNS2. Recent concert halls and opera house in Japan. Takayuki Hidaka (Takenaka R&D, 1-5-1 Otsuka Inzai, Chiba 270-1395, Japan, hidaka.takayuki@takenaka.co.jp)

Since we invited Dr. Beranek to Japan for the first time in 1989, we had been working together with him for a period of 13 years, until 2001, on seven hall projects as acoustic design consultants. All of these halls are of premium importance to Japan. Dr. Beranek always came up with innovative concepts and helped create halls endowed with high acoustic originality. These halls are now loved by music-related people and music fanciers and regarded as the pride of Japan. The reviews and studies achieved through these projects were published as seven J. Acoust Soc. Am papers to disclose the outcome in an objective way to the public. A brief outline of the history of our collaboration and its background are presented.

9:20

4aNS3. BBN and structural acoustics. Richard H. Lyon (RH Lyon Corp, 691 Concord Ave., Cambridge, MA 02138)

Leo Beranek, Dick Bolt, and Bob Newman founded a company based on technical excellence and client service—BBN. The early services were oriented to noise control and architectural acoustics, but these led fairly quickly over about a decade into several related fields. One such field, now called ''structural acoustics,'' arose from activities in noise control where the radiated sound due to vibrations of the machine caused problems. Commercial work on such problems was later augmented by work for the US Navy, the Air Force, and NASA. In the mid- and late-1950s Ira Dyer built the group that during the course of about the next decade developed the field of structural acoustics, with emphasis on statistical modeling and with applications to ships, aircraft, and space launch vehicles. The author will present some of his personal remembrances of this second decade, with particular emphasis on the development of statistical energy analysis.

9:40

4aNS4. Ocean acoustics at BBN and beyond. Ira Dyer (Dept. of Ocean Eng., MIT, 77 Massachusetts Ave., Cambridge, MA 02139, dyerira@msn.com)

Earlier known exclusively as underwater acoustics, ocean acoustics was pursued almost from BBN's beginning, the firm created by Richard Bolt and Leo Beranek. Ship noise, including that of submarines, and also sonar detection performance are perhaps still classified as underwater acoustics, while sound propagation and noise in the ocean have come to be classified as ocean acoustics. Given its creators, it is no surprise that contributions at BBN to ocean acoustics to about 1970 were based on fundamental concepts in architectural acoustics. These are in essence a set of geometrical ideas within which complicated multi-mode waves propagating in rooms can be reduced to simple theoretical and measurable aggregates expressing the mean-square wave motion. Thus, ray-averaged theories of propagation, wave coherence, and noise in the ocean were developed and used at BBN, as they were elsewhere, and are still in use today. By about 1970 it became clear that improved measurement tools, and increased understanding of the ocean state, including the bottom, could be exploited by theories and measurements of the phase as well as the mean-square motion. This story is traced from the perspective of those at BBN in the early years, to the present state of ocean acoustics.

10:00–10:20 Break

10:20

4aNS5. Noise policy: Past and present. William W. Lang (Noise Control Foundation, 29 Hornbeck Ridge, Poughkeepsie, NY 12603)

Leo Beranek and the firm Bolt Beranek and Newman have played a defining role in the formulation of noise policy in America. The firm that he and Richard Bolt founded in 1948 with fewer than a half-dozen others grew to become the world's largest acoustical consulting firm with more than 2000 employees. Two decades later in 1971, Leo Beranek was a key founder of the Institute of Noise Control Engineering of the U.S.A. The Institute, which inaugurated the INTER-NOISE series of annual noise congresses in 1972 under Beranek's direction, played a major role in the enactment by the Congress of the Noise Control Act of 1972. NCA-72 identified the Environmental Protection Agency as the leading Federal agency with oversight responsibilities responsible for implementing the noise policies defined by the Congress. In 1981, funding for EPA's noise program was withdrawn. Since then, leadership at the Federal level for implementing a coordinated national noise policy has been absent, but a dozen Federal agencies remain active in the noise field. With the exception of aircraft, no product emission regulations on major sources of noise are enforced today. To rectify this situation, Leo Beranek has recently been playing a leadership role in a concerted effort to rejuvenate America's national noise policy.

10:40

4aNS6. Standards for acoustical instruments. Alan H. Marsh (16072 Santa Barbara Ln., Huntington Beach, CA 92649-2155)

Prior to 1930, sounds were classified from subjective judgments of loudness. General-purpose instruments that could make objective measurements of sounds were not yet developed, nor were there any national or international standards with specifications for the performance characteristics of such instruments. In recognition of the need for instruments that could provide objective measurements of a sound, as well as the need for a national consensus on the characteristics of a sound to be measured, the newly founded Acoustical Society of America established Sectional Committees in 1930 to develop such standards under the procedures of the American Standards Association (now the American National Standards Institute). A draft of a standard for sound level meters prepared by the subcommittee for Sound Levels and Sound Level Meters was considered at a meeting of the Sectional Committee in

May 1934. After incorporating revisions in response to comments and suggestions, the draft standard for sound level meters was approved for publication by the American Standards Association on 17 February 1936. This paper describes the development of sound level meters and other instruments; along with related national and international standards, in the years since that auspicious beginning.

11:00

4aNS7. History of acoustical consulting at BBN. Eric W. Wood and Eric E. Ungar (Acentech, 33 Moulton St., Cambridge, MA 02138, ewood@acentech.com)

A book is being written describing the history of acoustical consulting at Bolt Beranek and Newman Inc. (BBN). It highlights people that joined the firm during the early years, the unique culture they established, the major clients and projects, noteworthy contributions made by the staff, and 40 years later the transition of BBNs acoustical consulting practice to Acentech. The book is based on interviews of colleagues and friends, primarily by Deborah Melone, a former BBN editor. This presentation reviews the early days of the consulting partnership formed by Bolt and Beranek in 1948, its initial growth, and its achievements.

11:20

4aNS8. Books on acoustics. Neil A. Shaw (Menlo Scientific Acoustics, Inc., P.O. Box 1610, Topanga, CA 90290, menlo@ieee.org)

The legacy of a man is not limited to just his projects. His writings in many cases are a more lasting, and a definitely more accessible, monument. For 60 years, Leo L. Beranek has produced books on acoustics, acoustic measurements, sound control, music and architecture, noise and vibration control, concert halls, and opera houses in addition to teaching and consulting. His books are standard references and still cited in other books and in technical and professional articles. Many of his books were among, if not, the first comprehensive modern treatment of the subject and many are still foremost. A review of Dr. Beranek's many books as well as some anecdotes about the circumstances and consequences of same will be presented.

11:35

4aNS9. My 65 years in acoustics. Leo L. Beranek (975 Memorial Dr., Ste. 804, Cambridge, MA 02138-5755, beranekleo@ieee.org)

My entry into acoustics began as research assistant to Professor F. V. Hunt at Harvard University. I received my doctorate in 1940 and directed the Electro-Acoustic Laboratory at Harvard from October 1940 until September 1945. In 1947, I became a tenured associate professor at MIT, and, with Richard H. Bolt, formed the consulting firm Bolt and Beranek, that later included Robert B. Newman, becoming BBN. My most significant contributions before 1970 were design of wedge-lined anechoic chambers, systemization of noise reduction in ventilation systems, design of the world's largest muffler for the testing of supersonic jet engines at NASA's Lewis Laboratory in Cleveland, speech interference level, NC noise criterion curves, heading New York Port Authority's noise study that resulted in mufflers on jet aircraft, and steep aircraft climb procedures, and publishing books titled, *Acoustical Measurements, Acoustics, Noise Reduction, Noise and Vibration Control*, and *Music, Acoustics and Architecture*. As President of BBN, I supervised the formation of the group that built and operated the ARPANET (1969), which, when split in two (using TCP/IP protocol) became the INTERNET (1984). Since then, I have written two books on *Concert Halls and Opera Houses* and have consulted on four concert halls and an opera house.

THURSDAY MORNING, 27 MAY 2004 CONFERENCE ROOM L, 9:00 TO 11:45 A.M.

Session 4aPA

Physical Acoustics: Guided Waves, Resonators and Structures

Sameer Madanshetty, Chair

Mechanical Engineering Department, Kansas State University, Rathbone Hall, Manhattan, Kansas 66506-5106

Contributed Papers

9:00

4aPA1. Optical measurement of acoustic streaming in a waveguide. Tetsushi Biwa (Nagoya Univ., Chikusa-ku, Nagoya 464-8603, Japan, biwa@nuap.nagoya-u.ac.jp), Michael W. Thompson, and Anthony A. Atchley (Penn State Univ., University Park, PA 16802)

The generation of time-averaged fluid velocities by acoustic waves is known as acoustic streaming. The recent development of traveling-wave thermoacoustic engines has accelerated the need for a deeper understanding of acoustic streaming. In these devices, mass flow associated with traveling-wave streaming results in a large heat leak. However, no quantitative measurements of traveling-wave streaming have been found in the literature. In the present work, we study acoustic streaming induced by traveling- and standing-wave fields in a cylindrical waveguide. Drivers attached to each end of the waveguide are driven at the same frequency. By tuning the relative magnitudes and phases of the voltages supplied to these drivers, both traveling- and standing-wave fields can be generated. Laser Doppler anemometry has been used to make quantitative measurements of acoustic streaming induced by a standing wave in a resonator [M. W. Thompson and A. A. Atchley, in *Nonlinear Acoustics at the Beginning of the 21st Century*, edited by O. V. Rudenko and O. A. Sapozhnikov (Moscow, 2002), Vol. 1, pp. 183–190]. We apply this experimental technique to the traveling-wave field in this waveguide, and compare the local acoustic variables and the streaming.

9:15

4aPA2. Optical measurement of acoustic streaming in a standing wave with a temperature gradient. Michael W. Thompson and Anthony A. Atchley (Penn State Grad. Prog. in Acoust., University Park, PA 16802, mwt126@psu.edu)

Laser Doppler anemometry (LDA) with burst spectrum analysis (BSA) is applied to the measurement of Rayleigh-type acoustic streaming in a cylindrical standing-wave resonator constructed partly of glass and filled with air. The resonator is driven sinusoidally, and the axial component of the streaming-velocity field is measured along the resonator's axis and across its diameter. At large acoustic amplitudes, distortion of the streaming field is observed. This distortion is attributed to two effects: (1) fluid inertia and (2) a steady temperature gradient resulting from thermoacoustic heat transport along the resonator's inner wall. Measurements of the streaming field are made at several different acoustic amplitudes and with the resonator either wrapped in foam insulation, suspended within an air-filled tank, or surrounded by a water jacket in order to control the magnitude of the temperature gradient.

9:30

4aPA3. Demonstration by optical visualization of acoustic re-radiation from A_0 and S_0 waves on submerged shells. P. K. Raju, A. C. Ahyi, H. Cao (Dept. of Mech. Eng., Auburn Univ., Auburn, AL 36849), and H. Überall (Catholic Univ., Washington, DC 20064)

Elastic waves propagating on thin shells may be classified like for the Lamb waves on plates (A_0, S_0, A_1, S_1 . . .), and the Scholte-Stoneley wave (A) in the fluid loading. The present study deals with evacuated shells of semi-infinite extent and a uniformly curved front surface, on which acoustic pulses are incident head-on. The incident signals cause the generation of the mentioned shell waves at a critical angle of incidence; these may be observed by their reradiation into the fluid at the same critical angle. Our demonstration of the reradiated pulses consists in a numerical evaluation of the incident and reradiated fields, and a visualization of the corresponding pulses in a tank experiment employing the Schlieren method. While the A wave could not be observed because of its rapid decay following its generation, we were able to demonstrate by both methods the generation of reradiated pulses of the A_0 and the S_0 wave, at the same time verifying the value of the critical angle of their generation.

9:45

4aPA4. Numerical simulation of flow-excited cavity tones. Paul J. Zoccola, Joseph F. Slomski, and Theodore M. Farabee (Naval Surface Warfare Ctr., Carderock Div., 9500 MacArthur Blvd., West Bethesda, MD 20817)

The objective was to develop a numerical model for investigating the effect of obstructions in the opening on the frequency and amplitude of flow-excited cavity pressure fluctuations. A hybrid Reynolds-averaged Navier–Stokes/large eddy simulation (RANS/LES) method was used. The RANS/LES method ensures that important turbulence scales outside of the boundary layer are resolved, while permitting more economical RANS modeling of the smaller turbulent scales immediately adjacent to walls. Solutions were obtained at the speed where the cavity Helmholtz resonance was excited and at a speed well below resonance. A solution was also obtained where the cavity opening was partially obstructed. Results were compared to experimental data obtained for the same configuration. In all cases, the frequency of the shear layer oscillations was accurately predicted. Resonant cases also showed excellent agreement on the amplitude of the first harmonic. Some initial difficulties in predicting the amplitude in the nonresonant case are attributed to excessive dissipation in the LES calculations. Nonresonant cases are being resimulated using a less dissipative advective scheme and SGS turbulence model. A comparison of some flow field results showed reasonable agreement in certain cases.

10:00–10:15 Break

10:15

4aPA5. Experimental and numerical investigation of flow through an oscillated acoustic resonator. Christopher C. Daniels (Univ. of Akron, Akron, OH 44325), Joshua Finkbeiner, Bruce M. Steinetz (NASA Glenn Res. Ctr., Cleveland, OH 44135), Mahesh Athavale, and Maciej Pindera (CFD Res. Corp., Huntsville, AL 35805)

An acoustic resonator was oscillated experimentally at the fundamental gas resonant frequency to develop standing pressure waves. The conical shaped resonator contained openings that provided an air passage from a pressurized cavity through the resonator to the ambient environment. For several pressure differentials applied across the resonator, the rate of air flow is reported for no resonator oscillation, and for on-resonant and near-resonant frequency oscillations. When compared to no oscillation and near-resonant frequency oscillation at all pressure differentials, the standing waves within the resonator reduced the flow of air through the system when oscillated on-resonance. A two-dimensional numerical model was developed using a commercial CFD package to simulate the gas flow within the system. The mass flow of air through the oscillating resonator was matched using the numerical simulations. For a low value of differential pressure, the simulations showed the reversal of gas flow into the high pressure cavity during a part of the cycle, while allowing free flow of air during other cycle phases. For a high value of differential pressure, no flow reversal was observed. The goal of the study was to determine the applicability of nonlinear acoustics to advanced seal concepts at NASA Glenn Research Center.

10:30

4aPA6. Vibroacoustics of three-dimensional drum silencer. Lixi Huang (Dept of Mech. Eng., The Hong Kong Polytechnic Univ., Kowloon, Hong Kong)

When low-frequency sound waves travel down a duct in which a segment of hard walls is replaced by membranes backed by side branch cavities, they are reflected as a result of sound radiation by the induced membrane vibration. The reflection is effective over a broad frequency band when high tension is applied in the axial direction of the duct. The device is thus called a drumlike silencer, and its existing vibroacoustics theory is based on a two-dimensional model in which the membrane behaves as a string. This study extends the theory to three dimensions, in which the membrane covers a finite rectangular cavity with all edges fixed. It is shown analytically and validated experimentally that the fixed edges of the drum silencer have no effect on the silencing performance of the device working in the cutoff frequency range of the duct and when there is no tension in the transverse direction. Finite-element method is used to simulate the full fluid–membrane coupling and the results are validated by known solutions. The validated numerical method is then applied to demonstrate the acoustic benefit of using a circular cavity for a rectangular duct.

10:45

4aPA7. Fluid-structure interaction in fast breeder reactors. A. A. Mitra, D. N. Manik (IIT Bombay, India), and P. A. Chellapandi (IGCAR, Kalpakkam, India)

A finite element model for the seismic analysis of a scaled down model of Fast breeder reactor (FBR) main vessel is proposed to be established. The reactor vessel, which is a large shell structure with a relatively thin wall, contains a large volume of sodium coolant. Therefore, the fluid structure interaction effects must be taken into account in the seismic design. As part of studying fluid-structure interaction, the fundamental frequency of vibration of a circular cylindrical shell partially filled with a liquid has been estimated using Rayleigh's method. The bulging and sloshing frequencies of the first four modes of the aforementioned system have been estimated using the Rayleigh–Ritz method. The finite element formulation of the axisymmetric fluid element with Fourier option (required due to seismic loading) is also presented.

11:00

4aPA8. Measurement of the frequency dependence of the ultrasonic parametric threshold amplitude. Alem A. Teklu (Dept. of Phys. and Astron., College of Charleston, 66 George St., Charleston, SC 29424, teklua@cofc.edu), Nico F. Declercq, and Mack A. Breazeale (The Univ. of Mississippi, Univ., MS 38677)

The excitation of a liquid-filled cavity by ultrasonic waves was described as a parametric phenomenon [L. Adler and M. A. Breazeale, J. Acoust. Soc. Am. **48**, 1077–1083 (1970)]. A standing ultrasonic wave is produced in the cavity by a drive transducer at one end and a rigid reflector at the other. Variations in the cavity length lead to frequencies lower than the drive frequency. Such a situation can be described by a modified Matthieu equation whose solution can be used to predict a threshold for parametric oscillation. Originally the threshold was assumed to decrease with frequency [L. Adler, Ph.D. dissertation, The University of Tennessee, 1969]. The apparatus used by Adler and Breazeale recently was refined for accurate measurement of the threshold amplitude for parametric excitation between 1 and 10 MHz. The measurement showed that in this range the threshold amplitude actually increases with increasing drive frequency. The results are compared with existing models.

11:15

4aPA9. Localized mode of sound in a waveguide with a pair of Helmholtz resonators. N. Sugimoto and H. Imahori (Dept. of Mech. Sci., Grad. School of Eng. Sci., Univ. of Osaka, Toyonaka, Osaka 560-8531, Japan)

This paper examines a localized mode of sound in a planar waveguide between rigid walls with a pair of identical Helmholtz resonators connected. Assuming the waveguide extends infinitely, a two-dimensional problem to the wave equation is solved within linear, lossless theory. By the localized mode are meant time-harmonic, stationary oscillations without radiation damping, which are confined only in the vicinity of the resonators and decay exponentially away from them. It is revealed that when the resonators are connected *vis-à-vis* to the upper and lower walls, an antisymmetric localized mode can exist but no symmetric mode exists. In the antisymmetric mode, the sound pressure changes out of phase π in the upper and lower half of the waveguide. In the dispersion diagram between frequency and wavenumber, the localized mode introduces an isolated frequency at vanishing wavenumber in addition to an infinite number of duct modes. The frequency of the localized mode is lower than the lowest cutoff frequency of the duct modes and also a natural frequency of the resonator. The solution for the localized mode is represented by superposition of an infinite number of antisymmetric, evanescent duct modes. When the resonators are staggered on both walls, the localized mode disappears.

11:30

4aPA10. Edge resonance in elastic bodies. The physical analysis of the phenomenon. Victor T. Grinchenko and Natalya S. Gorodetskaya (Inst. of Hydromechanics, Natl. Acad. of Sci. of Ukraine, Kiev 03680, Ukraine)

Edge resonance in elastic bodies of finite sizes is a specific phenomenon without analogs in wave fields in electrodynamics and acoustics. This phenomenon was examined by studying normal modes of finite elastic cylinders and plates. With excitation of the edge resonance, an unusual wave effect is formed. The eigenfrequency of an elastic body is independent of its geometrical size and, in some cases, even increases with increasing body dimension. Eigenfrequency of edge resonance is into the frequency range where only one propagating wave exists in the corresponding infinite cylinder or layer. It gives a basis for the general qualitative conclusion that edge resonance is a result of intensive excitation of evanescent waves. However, such an obvious idea does not open the nature of the edge mode formation. The analysis of longitudinal and shear components of propagating wave gives way to understanding specific features of reflection of normal wave in semi-infinite waveguides. The determining role of effects of transformation of longitudinal wave in shear wave and vice versa at reflection from free boundary is shown. Dependence of quantitative characteristics of these effects from Poisson's ratio allows estimation of the quality factor of edge resonance and changes corresponding eigenfrequency.

THURSDAY MORNING, 27 MAY 2004 ROYAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 4aPP

Psychological and Physiological Acoustics: Loudness and Binaural Perception

Mary Florentine, Cochair
Department of Speech–Language Pathology and Audiology, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115

H. Steven Colburn, Cochair
Department of Biomedical Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

8:00

4aPP1. A study of listening habits in adolescents: Correlating stated loudness preferences with actual listening levels. Laura Warren, Jean Warren (Columbia College Chicago, 1684 Linden St., Des Plaines, IL 60018, warr808@aol.com), and Dominique Cheenne (Columbia College Chicago, Des Plaines, IL 60018)

Evidence suggests that children are damaging their hearing in substantial numbers [Niskar *et al.*, J. Am. Med. Assoc. (1998)]. Conventional thinking would suggest that cultural norms and attitudes contribute to a desire in children to model what they have seen in the media, thus implying that they would be listening to music at levels that are considered harmful. Our study focused on a gender-balanced group of 316 elementary-age students and aimed at assessing a correlation between an attitudinal survey related to loud music and the children's own listening levels. The study was broader in scope and in sample size than previous work [Fucci, 138th ASA Meeting, 11/99]. Findings were both surprising and encouraging, citing that a majority of children who expressed favoritism towards loud music listened to the presented samples at lower levels than expected. The study also proposes a set of listening level distribution curves that may prove useful for future studies with older participants.

8:15

4aPP2. Extracting magnitude estimations of loudness from pairwise judgments. Eugene Galanter (Columbia Univ., 460 Riverside Dr., New York, NY 10027, eg53@columbia.edu)

Four problems limit widespread applications of magnitude estimation scales. There is first a serious question about the meaning of the judgments. Second, the judgments appear to be intrinsically unreliable. Third, amalgamation methods used to strike averages from several observers are poorly understood. Finally, it is difficult to know how to evaluate a single entity as distinct from a stimulus domain by these methods. A new psychophysical method is described and demonstrated that can scale any perceptual or attitudinal continuum. Observers give a numerical estimate of the magnitude of a randomly selected stimulus relative to another such stimulus to which a computer has assigned a random number. The stimulus pairs vary at random from trial to trial. Ratios of these computer-person number pairs estimate the slope of the psychophysical function. The slope lets us normalize the judgments which can then be mapped onto the stimulus domain. Loudness functions for individuals are shown with none of the cusps or singularities of traditional magnitude estimations from individuals. [Work supported in part by NASA.]

8:30

4aPP3. Directional loudness of narrow-band noises in an anechoic sound field. Ville P. Sivonen and Wolfgang Ellermeier (Dept. of Acoust., Aalborg Univ., Fredrik Bajers Vej 7 B5, DK-9220 Aalborg, Denmark, vps@acoustics.auc.dk)

In order to investigate the effect of sound incidence angle on loudness across a larger set of parameters than have been used in most previous studies, a listening experiment was carried out using a loudspeaker setup in an anechoic chamber. Eight subjects, whose absolute hearing thresholds and head-related transfer functions (HRTFs) were measured, participated in a total of 22 sessions each. On each trial their task was to judge which of two narrow-band noises sounded louder. These judgments were used in an adaptive procedure to find loudness matches between a frontal reference location and seven other sources, positioned both in the horizontal and median planes. Sound incidence angle, center frequency, and overall SPL were varied in the procedure. The results show that loudness is not constant over sound incidence angles, with matches varying over a range of 10 dB, and showing considerable frequency dependency. The pattern of results also varies substantially between subjects, but can be accounted for by interindividual variations in the listeners' HRTFs. [Work supported by Brüel & Kjær Sound & Vibration Measurement A/S.]

8:45

4aPP4. Perceptual and procedural learning in interaural cue discrimination. Helena Constantinides and David R. Moore (Univ. Lab. of Physiol., Parks Rd., Oxford OX1 3PT, UK, hconstantinides@doctors.org.uk)

Training in stimulus detection, discrimination, or identification generally leads to an improvement in performance, apparently in two phases. Learning during the early, rapid phase (within hours) has been attributed predominantly to procedural or task learning. Perceptual (true stimulus) learning is believed to occur mainly in the slower, later phase. Two experiments were undertaken to determine whether significant perceptual learning also occurs during the early phase. In the first, four groups of listeners were trained for 1 h in interaural level difference (ILD) discrimination. Stimuli were presented within a different, two- or three-interval task for each group. The following day, complete generalization of learning from the trained to the untrained tasks was demonstrated. In the second experiment, two groups of listeners were trained in an ILD discrimination task. In one group the stimulus was at a fixed, suprathreshold level (ILD = 15 dB), and in the other the stimulus was varied adaptively around the listener's discrimination threshold. The post-train mean stimulus discrimination threshold in the suprathreshold group was 3.9 dB, and in the threshold group was 2.4 dB ($p < 0.01$). These findings suggest perceptual learning contributes significantly to the early, rapid phase of performance improvement in interaural cue discrimination.

9:00

4aPP5. How do listeners detect interaural incoherence? Matthew J. Goupell and William M. Hartmann (Dept. of Phys. and Astron., Michigan State Univ., 4230 BPS Bldg., East Lansing, MI 48824)

Listeners are remarkably sensitive to interaural incoherence. They can easily distinguish between a coherence of 1.00 and a coherence of 0.99. Interaural incoherence leads to the sensation of apparent source width; it is also held to be responsible for the masking level difference and for the creation of binaural pitch. However, incoherence *per se* is only a statistical description of signals; it is not indicative of any particular auditory binaural property. To discover the relevant binaural attribute(s), three-interval oddity experiments were performed using narrow-band noises that were perfectly coherent or slightly incoherent. The listener's task was to identify the incoherent noise. First, it was shown that different frozen noises with identical values of incoherence could often differ greatly in detectability. Subsequent experiments studied the detectability of incoherent frozen noises that were were particularly strong or particularly weak according to selected binaural difference functions. These functions tested different rules for combining interaural phase differences (IPD) and interaural level differences (ILD)—either weighted sums of squares or rules that combine IPD and ILD to form a fluctuating lateralized image. The goal was to discover the best predictor of incoherence detectability. [Work supported by the NIDCD Grant DC 00181.]

9:15

4aPP6. Lateralization of the Huggins pitch. Peter Xinya Zhang and William M. Hartmann (Dept. of Phys. and Astron., Michigan State Univ., 4230 BPS Bldg., East Lansing, MI 48824)

The lateralization of the Huggins pitch (HP) was measured using a direct estimation method. The background noise was initially $N0$ or $N\pi$, and then the laterality of the entire stimulus was varied with a frequency-independent interaural delay, ranging from -1 to $+1$ ms. Two versions of the HP boundary region were used, stepped phase and linear phase. When presented in isolation, without the broadband background, the stepped boundary can be lateralized on its own but the linear boundary cannot. Nevertheless, the lateralizations of both forms of HP were found to be almost identical functions both of the interaural delay and of the boundary frequency over a two-octave range. In a third experiment, the same listeners lateralized sine tones in quiet as a function of interaural delay. Good agreement was found between lateralizations of the HP and of the corresponding sine tones. The lateralization judgments depended on the boundary frequency according to the expected hyperbolic law except when the frequency-independent delay was zero. For the latter case, the dependence on boundary frequency was much slower than hyperbolic. [Work supported by the NIDCD grant DC 00181.]

9:30

4aPP7. Earedness: Left-eared and right-eared listeners. William M. Hartmann, Peter Xinya Zhang (Dept. of Phys. and Astron., Michigan State Univ., 4230 BPS Bldg., East Lansing, MI 48824), and John F. Culling (Cardiff Univ., Cardiff CF1 3YG, UK)

The Huggins pitch (HP) stimulus known as HP- is created with a broadband background noise having an interaural phase difference of zero, together with a narrow boundary region wherein the interaural phase varies with frequency. At the spectral center of the boundary region the interaural phase is 180 deg. Therefore, HP- is symmetrical with respect to the two ears. Despite the symmetry, most listeners hear the HP image strongly lateralized to one side of the head. Some hear it on the right; others hear it on the left. Two surveys, involving 51 listeners, found that these perceptions do not change when the headphones are reversed. Extensive experiments with five listeners found that the lateralization directions were usually insensitive to variations in the frequency of the boundary region (more than two octaves). The left or right preference was strong enough that listeners chose alias locations (differing from a more central location by 360 deg) on the preferred side when various frequency-independent interaural delays and phase shifts were added to the HP

stimulus. The experiments suggest that given ambiguous stimuli, listeners exhibit earedness—a preference similar to, but not as strong as, handedness. [Work supported by the NIDCD Grant DC 00181.]

9:45

4aPP8. Auditory-evoked magnetic fields in relation to the interaural cross correlation. Yoshiharu Soeta, Seiji Nakagawa, Mitsuo Tonoike (Life Electron. Lab., Natl. Inst. of Adv. Industrial Sci. and Technol. (AIST), Midorigaoka, Ikeda, Osaka 563-8577, Japan, y.soeta@aist.go.jp), Takuya Hotehama, and Yoichi Ando (Kobe Univ., Rokkodai, Nada, Kobe 657-8501, Japan)

Noninvasive measurements utilizing magnetoencephalography (MEG) have been used to study how sound stimulus features are represented in the human brain. These measurements have successfully revealed how, for example, tone frequency, periodicity, and intensity are encoded. Here, the auditory-evoked magnetic fields in change of the magnitude of the interaural cross correlation (IACC) were analyzed. The IACC of the stimuli was controlled by mixing two independent bandpass noises in appropriate ratios. The auditory stimuli were binaurally delivered through silicon tubes and earpieces inserted into the ear canals. All source signals had the same sound-pressure level. Nine volunteers with normal hearing took part in this study. The auditory-evoked fields were recorded using a neuromagnetometer in a magnetically shielded room. Combinations of a reference stimulus (IACC=1.0) and test stimuli (IACC=0.2,0.6,0.85) were presented alternately at a constant 0.5-s interstimulus interval, and the MEGs were recorded and averaged more than 50 times. The results showed that the peak amplitude of N1m, which was found above left and right temporal lobes around 100 ms after the stimulus onset, significantly decreased with increasing IACC. The N1m latencies were not affected by IACC.

10:00–10:15 Break

10:15

4aPP9. Psychophysical calibration of auditory range control in binaural synthesis with independent adjustment of virtual source loudness. William L. Martens (Faculty of Music, McGill Univ., Montreal, QC H3A 1E3, Canada, wlm@music.mcgill.ca)

This paper reports the results of a study designed to evaluate the effectiveness of synthetic cues to the range of auditory images created via headphone display of virtual sound sources processed using individualized HRTFs. The particular focus of the study was to determine how well auditory range could be controlled when independent adjustment of loudness was also desired. Variation in perceived range of the resulting auditory spatial images was assessed using a two-alternative, forced choice procedure in which listeners indicated which of two successively presented sound sources seemed to be more closely positioned. The first of the two sources served as a fixed standard stimulus positioned using a binaural HRTF measured at ear level, 1.5 m from the listeners head at an azimuth angle of 120 deg. The second source served as a variable loudness comparison stimulus processed using the same pair of HRTFs, with the same interaural time difference but with a manipulated interaural level difference. From the obtained choice proportions for each pairwise comparison of stimuli, numerical scale values for auditory source range were generated using Thurstone's Case IV method for indirect scaling. Results provide a basis for calibrated control over auditory range for virtual sources varying in loudness.

10:30

4aPP10. Development of the sound localization cues in cats. Daniel J. Tollin (Dept. of Physiol., Univ. of Wisconsin–Madison, 1300 University Ave., Madison, WI 53706, tollin@physiology.wisc.edu)

Cats are a common model for developmental studies of the psychophysical and physiological mechanisms of sound localization. Yet, there are few studies on the development of the acoustical cues to location in cats. The magnitude of the three main cues, interaural differences in time (ITDs) and level (ILDs), and monaural spectral shape cues, vary with location in adults. However, the increasing interaural distance associated with a growing head and pinnae during development will result in cues that change continuously until maturation is complete. Here, we report measurements, in cats aged 1 week to adulthood, of the physical dimensions of the head and pinnae and the localization cues, computed from measurements of directional transfer functions. At 1 week, ILD depended little on azimuth for frequencies <6–7 kHz, maximum ITD was 175 μs, and for sources varying in elevation, a prominent spectral notch was located at higher frequencies than in the older cats. As cats develop, the spectral cues and the frequencies at which ILDs become substantial (>10 dB) shift to lower frequencies, and the maximum ITD increases to nearly 370 μs. Changes in the cues are correlated with the increasing size of the head and pinnae. [Work supported by NIDCD DC05122.]

10:45

4aPP11. Localization suppression and fusion measure of the precedence effect in young children. Ruth Litovsky, Shelly Godar, and Gongqiang Yu (Univ. of Wisconsin—Madison, Waisman Ctr., Madison, WI 53705)

This study investigated aspects of the precedence effect (PE) known as fusion and localization dominance in children 4–5 years of age. Stimuli were three, 25-ms noise bursts (2-ms rise/fall times) with 250-ms ISI. On PE conditions the lead stimulus was presented from one of six locations in azimuth, and the lag was at 0 deg. Lead-lag delays varied from 5 to 100 ms. Localization was measured using an identification paradigm. Fusion was measured separately whereby subjects reported whether a single auditory event or two auditory events were perceived. Children reported two sounds on 75% of trials (fusion threshold) at delays ranging from 15 to 35 ms. Below fusion thresholds, the localization of the lead was similar to that of single-source stimuli. Above fusion thresholds lead localization was significantly degraded, persisting out to 100 ms. Localization of the lag was poor at all delays on which it was reported as being heard. According to these results localization dominance (difficulty localizing the lag) in children persists at greater delays than fusion, which is consistent with findings obtained in adult subjects. The range of delays over which these effects are robust in children is longer than the range observed in adults.

11:00

4aPP12. An evaluation of two frequency-based mechanisms for direct-to-reverberant energy ratio discrimination. Nandini Iyer, Erik Larsen, Charissa Lansing, and Albert Feng (Beckman Inst. for Adv. Sci. and Technol., Univ. of Illinois at Urbana–Champaign, 405 N. Mathews Ave., Urbana, IL 61801)

To determine the mechanism by which direct-to-reverberant energy ratio (D/R) is discriminated, a previous experiment (Larsen *et al.*, 2003) reported JNDs at three different D/R values (-10, 0, $+10$ dB). Of three proposed mechanisms, the data favored a model based on detection of spectral variation of the magnitude of the source-receiver transfer function. However, a model based on detection of spectral centroid of the received signal could not be rejected conclusively. In order to determine the relative salience of spectral variance versus spectral centroid cues, the current study reports findings from two experiments. These data will be useful in the development of a general model for distance perception. D/R was fixed at 0 dB, because D/R discrimination is most acute (Larsen *et al.*, 2003) and spectral variations are largest (Jetzt, 1979) at that ratio. In experiment one, signals (300 ms Gaussian noise; 10 ms onset/offset time) were modified such that spectral variation cues were removed by roving the compression of the spectral envelope variations. In the second experiment, spectral centroid cues were removed by using band-limited signals (300 ms Gaussian noise; 10 ms onset/offset time) and roving the center frequency. Listener sensitivity in both these experiments will be compared to model predictions.

11:15

4aPP13. The influence of target-masker similarity on across-ear interference in dichotic listening. Douglas Brungart and Brian Simpson (Air Force Res. Lab., 2610 Seventh St., Wright-Patterson AFB, OH 45433)

In most dichotic listening tasks, the comprehension of a target speech signal presented in one ear is unaffected by the presence of irrelevant speech in the opposite ear. However, recent results have shown that contralaterally presented interfering speech signals do influence performance when a second interfering speech signal is present in the same ear as the target speech. In this experiment, we examined the influence of target-masker similarity on this effect by presenting ipsilateral and contralateral masking phrases spoken by the same talker, a different same-sex talker, or a different-sex talker than the one used to generate the target speech. The results show that contralateral target-masker similarity has the greatest influence on performance when an easily segregated different-sex masker is presented in the target ear, and the least influence when a difficult-to-segregate same-talker masker is presented in the target ear. These results indicate that across-ear interference in dichotic listening is not directly related to the difficulty of the segregation task in the target ear, and suggest that contralateral maskers are least likely to interfere with dichotic speech perception when the same general strategy could be used to segregate the target from the masking voices in the ipsilateral and contralateral ears.

11:30

4aPP14. The relative immunity of high-frequency transposed stimuli to low-frequency binaural interference. Leslie R. Bernstein and Constantine Trahiotis (Dept. of Neurosci. and Dept. of Surgery (Otolaryngol.), Univ. of Connecticut Health Ctr., Farmington, CT 06030)

We have recently demonstrated that high-frequency transposed stimuli, having envelopes designed to provide high-frequency channels with information similar to that normally available in only low-frequency channels, yield threshold-ITDs and extents of laterality comparable to those obtained with conventional low-frequency stimuli. This enhanced potency of ITDs conveyed by high-frequency transposed stimuli, as compared to conventional high-frequency stimuli, suggested to us that ITDs conveyed by transposed stimuli might be relatively immune to the presence of low-frequency binaural interferers. To investigate this issue, threshold-ITDs and extents of laterality were measured with a variety of conventional and transposed targets centered at 4 kHz. The targets were presented either in the presence or absence of a simultaneously gated diotic noise centered at 500 Hz, the interferer. As expected, the presence of the low-frequency interferer resulted in substantially elevated threshold-ITDs and reduced extents of laterality for the conventional high-frequency stimuli. In contrast, these interference effects were either greatly attenuated or absent for ITDs conveyed by the high-frequency transposed targets. The results will be discussed in the context of current models of binaural interference. [Work supported by NIH DC 04147, NIH DC04073, NIH DC 002304.]

11:45

4aPP15. Performance benefits of adaptive, multimicrophone, interference-canceling systems in everyday environments. Joseph G. Desloge, Martin J. Zimmer, and Patrick M. Zurek (Sensimetrics Corp., 48 Grove St., Somerville, MA 02144, desloge@sens.com)

Adaptive multimicrophone systems are currently used for a variety of noise-cancellation applications (such as hearing aids) to preserve signals arriving from a particular (target) direction while canceling other (jammer) signals in the environment. Although the performance of these systems is known to degrade with increasing reverberation, there are few measurements of adaptive performance in everyday reverberant environments. In this study, adaptive performance was compared to that of a simple, non-adaptive cardioid microphone to determine a measure of adaptive benefit. Both systems used recordings (at an *Fs* of 22 050 Hz) from the same two omnidirectional microphones, which were separated by 1 cm. Four classes of environment were considered: outdoors, household, parking garage, and public establishment. Sources were either environmental noises (e.g., household appliances, restaurant noise) or a controlled noise source. In all situations, no target was present (i.e., all signals were jammers) to obtain maximal jammer cancellation. Adaptive processing was based upon the Griffiths–Jim generalized sidelobe canceller using filter lengths up to 400 points. Average intelligibility-weighted adaptive benefit levels at a source distance of 1 m were, at most, 1.5 dB for public establishments, 2 dB for household rooms and the parking garage, and 3 dB outdoors. [Work supported by NIOSH.]

THURSDAY MORNING, 27 MAY 2004 LIBERTY 1/2, 8:15 TO 11:45 A.M.

Session 4aSA

Structural Acoustics and Vibration: Fluid/Structural Interaction: Radiation, Scattering and Transmission

Jeffrey S. Vipperman, Chair

Department of Mechanical Engineering, University of Pittsburgh, 648 Bennedum Hall, 3700 Ohara, Pittsburgh, Pennsylvania 15261

Contributed Papers

8:15

4aSA1. Transmission loss of a submerged viscoelastic slab containing oblate spheroidal coated microinclusions. Michael Haberman (Georgia Tech Lorraine, 2 rue Marconi, 57070-Metz, France), Yves Berthelot (Georgia Inst. of Technol., Atlanta, GA 30332-0405, yves.berthelot@me.gatech.edu), and Mohammed Cherkaoui (Universite de Metz, 57045 Metz, France)

The three-phase self-consistent (SC) micromechanical model of Cherkaoui *et al.* [J. Eng. Mater. and Technol. **116**, 274–278 (1994)] is employed to model the anisotropic effects induced by identically oriented coated ellipsoidal microinclusions embedded in a viscoelastic matrix. Numerical predictions of the lossy and anisotropic behavior of the viscoelastic composite material are obtained. This study is an extension of a previous paper [J. Acoust. Soc. Am. **112** (5), 1937–1943 (2002)], where the agreement of the micromechanical approach with the long wavelength scattering model of Baird *et al.* [J. Acoust. Soc. Am. **105** (3), 1527–1538 (1999)] was demonstrated. The use of the SC model for the special case of a biphase effective material is also considered and compared with the complex bounds given by Gibiansky and Lakes [J. Mech. Mater. **16**, 317–331 (1993)]. Parametric studies are presented for the transmission loss in the 0–100-kHz frequency range of a 1-cm-thick composite material slab containing oblate spheroidal inclusions of varying aspect ratios when the slab is submerged under water. The results are compared with the case of a composite material having identical constituent material properties and volume fractions as the oblate case but containing spherical inclusions.

8:30

4aSA2. The transmission loss through curved sandwich composite structures. Sebastian Ghinet, Noureddine Atalla (Dept. of Mech. Eng., Universite de Sherbrooke, 2500 Blvd. Universite, Sherbrooke, QC J1K 2R1, Canada), and Haisam Osman (The Boeing Co., Huntington Beach, CA 92647)

The principal aim of this work is to present a model for the transmission loss of sandwich composite cylindrical shells. The effects of membrane, bending, and transverse shearing as well as rotational inertia are considered in all of the layers composing the structure. The elastic constants of any layer are related to the orthotropic angle-ply defined as the angle of the principal directions of the layers material to the global axis of the shell. Fundamental relations are expressed using the dynamic equilibrium relations of the unit forces in the structure. The structural impedance, critical frequencies and ring frequency are computed numerically in the general case of symmetrical laminated composite shell. Their expressions are developed in a wave approach context. A general eigenvalue approach to compute the dispersion curves of such structures is presented. Using these curves, the radiation efficiency, the modal density, the group velocity an the resonant and nonresonant transmission loss are computed and used within SEA framework to predict the sound transmission loss of these structures. Comparisons with existing models and experimental data are also discussed.

8:45

4aSA3. Theoretical investigation of noise transmission into a finite cylinder. Deyu Li and Jeffrey S. Vipperman (Dept. of Mech. Eng., Univ. of Pittsburgh, Pittsburgh, PA 15261)

A new mathematical model for characterizing noise transmission into a finite elastic cylindrical structure with application to a ChamberCore composite cylinder is presented. A plane wave obliquely impinges on the structure, the external sound field is approximated by the solution for an infinite cylinder, and the internal sound field is solved with the structural and acoustic modal interaction method. The noise reduction spectrum for characterizing noise transmission into the cylinder is defined, and the analytical model for the calculation of the noise reduction spectrum is developed. The analytical results show that the cavity resonances dominate the noise transmission into the finite cylinder, and the longitudinal acoustic modes play an important role in the noise transmission at the low frequencies. These results are matched with experimental results.

9:00

4aSA4. Noise control of submerged sources using close fitting enclosures. Joseph M. Cuschieri (Lockheed Martin Corp., MS2 Undersea Systsems, Riviera Beach, FL 33404)

Controlling the sound radiation from a source can be either performed at the source or in the transmission path or both. Modifications of the transmission path can be in the form of a close fitting enclosure. Simple models are available to estimate the effectiveness of the close fitting enclosure. However, if the sound source is submerged in water, the presence of the fluid within and external to the close fitting enclosure, and the type of absorption material within the enclosure, may influence the enclosure effectiveness. Modeling the fluid solely by its characteristic impedance may not lead to reasonable results. The influence of the fluid on the behavior of the close fitting enclosure should be taken into account in the modeling. Numerical results are presented that compare the insertion loss estimated using ''simple'' close fitting enclosure models to that from multi-layer theory. [Work supported by ONR.]

9:15

4aSA5. The general 2-D moments via integral transform method for acoustic radiation and scattering. Jerry R. Smith, Jr. (Naval Surface Warfare Ctr., Carderock Div., 9500 MacArthur Blvd., West Bethesda, MD 20817, smithjr@nswccd.navy.mil) and Mark S. Mirotznik (The Catholic Univ. of America, Washington, DC 20064)

The moments via integral transform method (MITM) is a technique to analytically reduce the 2-D method of moments (MoM) impedance double integrals into single integrals. By using a special integral representation of the Green's function, the impedance integral can be analytically simplified to a single integral in terms of transformed shape and weight functions. The reduced expression requires fewer computations and reduces the fill times of the MoM impedance matrix. Furthermore, the resulting integral is analytic for nearly arbitrary shape and weight function sets. The MITM technique is developed for mixed boundary conditions and predictions with basic shape and weight function sets are presented. Comparisons of accuracy and speed between MITM and brute force are presented. [Work sponsored by ONR and NSWCCD ILIR Board.]

9:30

4aSA6. Improved theory for acoustic radiation from a stiffener attached to a plate. Robert C. Haberman (BBN Technologies, Old Mystic Mill, 11 Main St., Mystic, CT 06355, rhaberman@bbn.com)

A flexural wave in a plate that interacts with a stiffener produces sound. Physically the stiffener provides concentrated line forces and moment onto the plate that generates a subset of waves with wave numbers less than the acoustic wave number. Typically the stiffener is modeled as a beam containing bending and torsional waves and mathematically coupled to the flexural waves in the plate. In reality, however, the stiffener exhibits resonant modes and is coupled to in-plane compression and shear waves in the plate. These effects produce in-plane forces at the plate–stiffener line junction (i.e., shear tractions), and highly frequency-dependent interaction forces and torques that result in complex frequency-dependent sound waves along with scattered and transmitted waves in the plate and stiffener. The theory associated with resonant stiffeners fully coupled to all three waves types in a plate with heavy fluid loading is presented, along with several numerical examples. The examples will demonstrate the conditions under which the simpler models are valid and conditions that require more sophisticated modeling. Application of the theory to problems in statistical energy analysis is briefly discussed.

9:45

4aSA7. Prediction of the sound reduction index with the modal theory. Alain Tisseyre, Cecile Courne, Andre Moulinier, and Thomas Buzzi (Tisseyre & Associes, 16 chemin de Manel, 31 400 Toulouse, France)

A procedure of sound reduction index calculation was developed on the basis of the modal theory. This calculation uses the study of the vibratory behavior and of the radiation of the plate. The vibratory study makes it possible to characterize the displacement of the wall. The latter is developed over a base: sinusoidal or polynomial. The use of a sinusoidal base makes it possible to obtain satisfactory results without limitation of frequency. In this case, the only possible boundary condition corresponds to a plate simply supported. The use of a polynomial base only makes it possible to obtain results in the low frequencies. On the other hand, in this case, several boundary conditions are possible: embedded, simply supported. In both cases, modeling several types of walls is possible: isotropic or orthotropic wall, single or multiple walls. Walls, orthotropic due to their geometry, were more particularly studied. It is the case in particular of alveolar slabs, bricks and ribbed sheets. A formulation of the acoustic transparency was established and programmed, and various results are presented. Currently, a calculation algorithm is developed in order to allow polynomial base modeling on a more significant frequency range.

10:00–10:15 Break

10:15

4aSA8. Numerical solution of acoustic radiation from fluid-loaded periodic structures using local-global homogenization. Pavel V. Danilov and Donald B. Bliss (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708)

Fluid-loaded vehicle structures, such as fuselages and hulls, often have spatially periodic discontinuities such as braces, ribs, and attachments. The structural motion, the acoustic radiation and scattering, and the interior sound field are of interest. Calculating the motion of fluid-loaded structures is a difficult task because of the high complexity and a disparity of length scales requiring high numerical resolution. Discontinuities cause the structural response to occur in a broad spectrum of spatial wavenumbers, and to exhibit stop-band and pass-band behavior. Structural discontinuities broaden the spatial wavenumber spectrum, causing both supersonic (radiating) and subsonic (nonradiating) waves. An analysis method called local-global homogenization (LGH) is used to predict directly the low wavenumber smooth response of periodic fluid-loaded structures in a self-contained manner. The low wavenumber part of the response is efficiently coupled to the acoustic field, since low wavenumbers correspond to supersonic phase speeds. In the LGH reformulation, an infinite order operator that must be truncated for numerical solution governs the equivalent smooth global problem. The numerical implementation is described, including the treatment of boundary conditions, and sample calculations are compared to exact solutions. The size of the computational problem is dramatically reduced by the LGH analytical reformulation.

10:30

4aSA9. A broadband energy-based boundary element method for predicting vehicle interior noise. Linda Franzoni, Jerry Rouse, and Tracy DuVall (Dept. of Mech. Eng., Duke Univ., Durham, NC 27708-0300)

Efficient calculation of vehicle interior noise is a challenging task. Classical acoustic boundary element calculations become costly at high frequencies due to the very large number of elements required and must be solved repeatedly for broadband applications. An alternative energy-intensity boundary element method has been formally developed that employs uncorrelated broadband directional intensity sources to predict mean-square pressure distributions in enclosures. The boundary source directivity accounts for local correlation effects and specular reflection. The method is applicable to high modal density fields, but it is not restricted to the usual low-absorption, diffuse, and quasi-uniform assumptions. The approach can accommodate fully specular reflection, or any combination of diffuse and specular reflection. This new method differs from the classical version in that the element size is large compared to an acoustic wavelength and equations are not solved on a frequency-by-frequency basis. These differences lead to an orders-of-magnitude improvement in computational efficiency. In vehicle interiors the sources are typically the vibrating walls of the enclosure. A special treatment for wall vibration sources has been developed for use with the new boundary element method. Calculations of spatially varying mean-square pressures agree well with computationally intensive modal solutions.

10:45

4aSA10. A viscous-acoustic boundary element formulation based on analytical-numerical matching. Donald B. Bliss and Matthew S. Swingle (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708)

Analytical-numerical matching (ANM) is an analysis scheme that combines a low-resolution global numerical solution with a high-resolution local analytical solution and a matching solution to form a uniformly valid composite solution. The application of ANM to harmonically oscillating bodies in a fluid leads to a novel reformulation of boundary element problems in acoustics and aeroacoustics. The singular kernal encountered in the integral equation is handled simply within the analytical local solution. The numerical implementation utilizes a smoothed Green's function solution to the governing equation with a distributed source term. Because the singular behavior has been removed from the numerical aspect of the problem, the approach converges rapidly and exhibits insensitivity to node (control point) location. The method allows low-resolution numerics to be combined with analytical corrections to obtain high accuracy solutions using a robust calculation scheme. The method has recently been extended to include viscous effects within the local solution. The appropriately modified global solution remains irrotational and can be expressed in terms of a smoothed potential. Sample results are shown for radiation from plates up to high frequencies. Ongoing research on ANM BEM is described, including work to include nonlinear convection within the viscous flow effects.

11:00

4aSA11. Acoustic emission analysis of shuttle thermal protection system. John Lane, Jeffery Hooker, Christopher Immer (ASRC Aerosp., M/S: ASRC-10, KSC, FL 32899), and James Walker (NASA, Huntsville, AL 35812)

Acoustic emission (AE) signals generated from projectile impacts on reinforced and advanced carbon/carbon (RCC and ACC) panels, fired from a compressed-gas gun, identify the type and severity of damage sustained by the target. This type of testing is vital in providing the required return to flight (RTF) data needed to ensure continued and safe operation of NASAs Space Shuttle fleet. Conventional AE analysis techniques require time domain processing of impulse data, along with amplitude distribution analysis. It is well known that identical source excitations can produce a wide range of AE signal amplitudes. In order to satisfy RTF goals, it is necessary to identify impact energy levels above and below damage thresholds. Spectral analysis techniques involving joint time frequency analysis (JTFA) are used to reinforce time domain AE analysis. JTFA analysis of the AE signals consists of short-time Fourier transforms (STFT) and the Huang–Hilbert transform (HHT). The HHT provides a very good measure of the instantaneous frequency of impulse events dominated by a single component. Identifying failure modes and cracking of fibers from flexural and/or extensional mode acoustic signals will help support in-flight as well as postflight impact analysis.

11:15

4aSA12. The elastic shell T-matrix theory evaluated. Michael Werby (NRL Code 7181, Stennis Space Ctr., MS 39529) and H. Uberall (Catholic Univ. of America, Washington, DC 20064)

The T-matrix method initiated by Waterman was extended in a clever paper by Peterson and the Varadans in 1980 to the complicated problem of elongated elastic shells. There have been some questions on whether the method converges for thin shells due to the impression that one must be able to inscribe a spherical surface in the annular region of the shell. The method may be derived by making use of several of the constraining equations that arise naturally and we show that the expression derived in 1980 is generally correct at least for objects with mirror symmetry. We present some details of the theoretical development with some calculations.

11:30

4aSA13. Reciprocity in the wave reflection and transmission problem. Yuri I. Bobrovnitskii (Dept. of Vibroacoustics, Mech. Eng. Res. Inst., 4 Maly Kharitonievsky Str., 101990 Moscow, Russia, bobrovni@orc.ru)

Matrices of reflection and transmission coefficients of plane waves in media or normal waves in waveguides are, in general case, not symmetric. When all the waves are of the propagating type, the matrix can be symmetrized by normalizing the wave amplitudes with the wave power flow. For evanescent (inhomogeneous) waves this is not valid because of their zero power flow. This difficulty is overcome in the present paper. The main result is that a reflection and transmission matrix becomes symmetric if the wave amplitudes are normalized with the certain energy-like quan-

tity that coincides with the power flow in the case of propagating waves. The result is valid for all types of waves (including those with complex wave numbers) and for all media and waveguides where the classical reciprocity theorem is valid. All symmetry relations, known in the literature for reflection and transmission coefficients follow from the result as particular cases. The result is useful in analysis of multimodal sound fields in composite media. It is illustrated in examples with inhomogeneous waves in fluids and solid waveguides.

THURSDAY MORNING, 27 MAY 2004 IMPERIAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 4aSC

Speech Communication: Poster Session III

Melissa Epstein, Chair
Biomedical Sciences, University of Maryland Dental School, 666 West Baltimore Street, Baltimore, Maryland 21201

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

4aSC1. Aeroacoustics of [s]. Michael S. Howe (College of Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, mshowe@bu.edu) and Richard S. McGowan (CReSS LLC, Lexington, MA 02420)

The theory of the sibilant fricative [s] is formulated and solved as a mathematical problem of aeroacoustics. Air is forced through the constriction between the tongue and the hard palate by the intra-oral pressure, forming a jet that strikes the upper incisors and leaves the mouth through a gap between the upper and lower incisors. The principal source of sound is the diffraction of jet turbulence pressure fluctuations by the incisors. The spectrum of these pressure fluctuations incident on the teeth is modeled analytically using an empirical formula adapted from boundary layer theory. Predictions are made of the far field acoustic pressure spectrum by reference to measured and estimated values of vocal tract dimensions and intra-oral pressure. Predicted spectra compare well with observations. The principal spectral peaks are determined by vocal tract physiology anterior to the tongue–palate constriction. The theory furnishes the first correct predictions of the dependence of the overall sound pressure level on the intra-oral pressure. This presentation will interpret the mathematical model in a nonmathematical manner. [Work supported, in part, by Grant NIDCD-01247 to CReSS LLC.]

4aSC2. The stop voicing contrast in French: From citation speech to sentencial speech. Nassima Abdelli-Beruh (Dept. of Speech-Lang. Pathol. and Audiol., New York Univ., 719 Broadway, Ste. 200, New York, NY 10003 and Dept. of Speech and Hearing Sci., City Univ. of New York, 365 Fifth Ave., New York, NY 10016), Eileen DeMaio (New York Univ., New York, NY 10003), and Miwako Hisagi (City Univ. of New York, New York, NY 10016)

This study explores the influence of speaking style on the salience of the acoustic correlates to the stop voicing distinction in French. Monolingual French speakers produced twenty-one **C̲**v**C̲** syllables in citation speech, in minimal pairs and in sentence-length utterances (/pa/_/a/ context: /il a di pa **C̲**v**C̲** a lui/; /pas/_/s/ context: /il a di pas **C̲**v**C̲** sa̲ lui/). Prominent stress was on the **C̲**v**C̲**. Voicing–related differences in percentages of closure voicing, durations of aspiration, closure, and vowel were analyzed as a function of these three speaking styles. Results show that the salience of the acoustic–phonetic segments present when the syllables are uttered in isolation or in minimal pairs is different than when the syllables are spoken in a sentence. These results are in agreement with findings in English.

4aSC3. Contextual sensitivity of some acoustic correlates to the stop voicing distinction in French. Nassima Abdelli-Beruh (Dept. of Speech-Lang. Pathol. and Audiol., New York Univ., 719 Broadway, Ste. 200, New York, NY 10003 and Dept. of Speech and Hearing Sci., City Univ. of New York, 365 Fifth Ave., New York, NY 10016)

This study examined how monolingual French speakers produced the stop voicing distinction in syllable-initial and syllable final stops embedded in various sentence contexts. Voicing-related differences in percentages of closure voicing, durations of aspiration, closure, and vowel were analyzed as a function of two experimental variables: the voicing class of the sound adjacent to the target stop [voiced vowel (/pa/_/a/ context), voiceless consonant (/pas/_/s/ context)] and the position of the stop within a syllable (syllable-initial, -final). Results from ANOVA showed that despite variations among speakers, group-based patterns surfaced in three contexts (i.e., syllable-initial stops in the /pa/_/a/ and /pas/_/s/ contexts and syllable-flnal stops in the /pa/_/a/ context): /b, d, g/ were more aspirated, preceded by longer vowels and were more frequently phonated than /p, t, k/. Closure durations for /b, d, g/ were shorter than those for /p, t, k/ in the /pa/-/a/ context only. Group patterns were not found for syllable-final stops in the /pas/-/s/ context. Results from discriminant analyses indicated that closure voicing was the variable that contributed the most to the phonological voicing distinction in all conditions.

4aSC4. The effect of emphatic stress on CV coarticulation. Golnaz Modarresi (Dept. of Linguist., B5100, Univ. of Texas, 1 University Station, Austin, TX 78712-0198, golie@mail.utexas.edu) and Harvey M. Sussman (Univ. of Texas, Austin, TX 78712-0198)

The effect of emphatic stress on CV coarticulation was investigated in the speech of one male and one female native speaker of American English using locus equation slope as a measure of CV coarticulation. Stressed real word $C_1V_2C_2$ tokens where C_1=/b,d,g/ and V_2=/i, I, e, ɛ, æ, u, o, ɔ, a/ were put in carrier sentences with the, thirty, or two preceding the test word. Each sentence was read three times in a normal manner and three times with emphasis on the test token. This resulted in a total of 486

tokens per speaker (3 stop consonants * 3 V_1 contexts * 9 V_2 contexts *2 emphasis patterns *3 repetitions). Locus equation slopes were derived by plotting $F2$ onset of C_1 against V_2 $F2$ mid-vowel frequency and fitting a regression line to data points. Consonant closure duration, V_2 duration, F_0, and amplitude were also measured. Despite a significant increase in the acoustic correlates of emphasis, locus equation slopes remained constant as a function of emphasis and varied as a function of place of articulation. This study provides further evidence of the stability of locus equation slopes as phonetic descriptors of stop place of articulation. [Work supported by NIH.]

4aSC5. The stability of locus equation slopes across stop consonant voicing/aspiration. Harvey M. Sussman (Dept. of Linguist., B5100, Univ. of Texas, 1 University Station, Austin, TX 78712-0198, sussman@mail.utexas.edu) and Golnaz Modarresi (Univ. of Texas, Austin, TX 78712-0198)

The consistency of locus equation slopes as phonetic descriptors of stop place in CV sequences across voiced and voiceless aspirated stops was explored in the speech of five male speakers of American English and two male speakers of Persian. Using traditional locus equation measurement sites for $F2$ onsets, voiceless labial and coronal stops had significantly lower locus equation slopes relative to their voiced counterparts, whereas velars failed to show voicing differences. When locus equations were derived using $F2$ onsets for voiced stops that were measured closer to the stop release burst, comparable to the protocol for measuring voiceless aspirated stops, no significant effects of voicing/aspiration on locus equation slopes were observed. This methodological factor, rather than an underlying phonetic-based explanation, provides a reasonable account for the observed flatter locus equation slopes of voiceless labial and coronal stops relative to voiced cognates reported in previous studies [Molis *et al.*, J. Acoust. Soc. Am. **95**, 2925 (1994); O. Engstrand and B. Lindblom, PHONUM **4**, 101–104]. [Work supported by NIH.]

4aSC6. The Nationwide Speech Project: A multi-talker multi-dialect speech corpus. Cynthia G. Clopper and David B. Pisoni (Speech Res. Lab., Dept. of Psych., Indiana Univ., Bloomington, IN 47405, cclopper@indiana.edu)

Most research on regional phonological variation relies on field recordings of interview speech. Recent research on the perception of dialect variation by naive listeners, however, has relied on read sentence materials in order to control for phonological and lexical content and syntax. The Nationwide Speech Project corpus was designed to obtain a large amount of speech from a number of talkers representing different regional varieties of American English. Five male and five female talkers from each of six different dialect regions in the United States were recorded reading isolated words, sentences, and passages, and in conversations with the experimenter. The talkers ranged in age from 18 and 25 years old and they were all monolingual native speakers of American English. They had lived their entire life in one dialect region and both of their parents were raised in the same region. Results of an acoustic analysis of the vowel spaces of the talkers included in the Nationwide Speech Project will be presented. [Work supported by NIH.]

4aSC7. Effects of coda voicing and aspiration on Hindi vowels. Claire Lampp and Heidi Reklis (Dept. of Linguist., Univ. of North Carolina at Chapel Hill, CB #3155, Chapel Hill, NC 27599, lampp@email.unc.edu)

This study reexamines the well-attested coda voicing effect on vowel duration [Chen, Phonetica **22**, 125–159 (1970)], in conjunction with the relationship between vowel duration and aspiration of codas. The first step was to replicate the results of Maddieson and Gandour [UCLA Working Papers Phonetics **31**, 46–52 (1976)] with a larger, language-specific data set. Four nonsense syllables ending in [open-o] followed by [k, kh, g, gh] were read aloud in ten different carrier sentences by four native speakers of Hindi. Results confirm that longer vowels precede voiced word-final consonants and aspirated word-final consonants. Thus, among the syllables, vowel duration would be longest when preceding the voiced aspirate [gh]. Coda voicing, and thus, vowel duration, have been shown to correlate negatively to vowel $F1$ in English and Arabic [Wolf, J. Phonetics **6**, 299–309 (1978); de Jong and Zawaydeh *ibid*, **30**, 53–75 (2002)]. It is not known whether vowel $F1$ depends directly on coda voicing, or is determined indirectly via duration. Since voicing and aspiration both increase duration, $F1$ measurements of this data set (which will be presented) may answer that question.

4aSC8. The effect of consonantal context on intensity distribution in vowels. Ewa Jacewicz and Robert Allen Fox (Dept. of Speech and Hearing Sci., The Ohio State Univ., Columbus, OH 43210, jacewicz.1@osu.edu)

Overall vowel intensity varies intrinsically with vowel quality and in terms of immediate consonantal context. These variations affect the levels of formants $F1$–$F4$ and $F0$, which change in consonantal contexts according to a vowel−specific pattern [Jacewicz, J. Acoust. Soc. Am. **114**, 2395 (2003)]. The current study further examines the differential distribution of acoustic energy across and within continuous frequency bands in the vowel as a function of vowel quality and consonantal context. The analysis technique used here was adopted from Sluijter and Van Heuven [J. Acoust. Soc. Am. **100**, 2471−2485 (1996)], who measured differences in spectral intensity distribution across frequency bands as a function of stress. The intensity distribution of the vowels /i, ɪ, ɛ, æ, u, ʊ/ in American English were measured across eight frequency bands in a stressed $[C1VC2]$ context (where $C1 = C2$). Contexts were chosen which have the greatest (e.g., voiceless fricatives) and the smallest (e.g., voiced stops) effect on changes in relative amplitude of formant frequencies and $F0$ for each particular vowel. Preliminary results indicate that intensity variations across frequency bands as a function of consonantal context have a potential to change the overall spectral balance. [Work supported by NIDCD R03 DC005560−01.]

4aSC9. Subglottal coupling and vowel space. Xuemin Chi and Morgan Sonderegger (Speech Commun. Group, RLE, MIT, Cambridge, MA 02139, smore@mit.edu)

A model of acoustic coupling between the oral and subglottal cavities predicts discontinuities in vowel formant prominences near resonances of the subglottal system. One discontinuity occurs near 1300–1500 Hz, suggesting the hypothesis that this is a quantal effect [K. N. Stevens, J. Phonetics **17**, 3–46 (1989)] dividing speakers' front and back vowels. Recordings of English vowels (in /hVd/ environments) for several male and female speakers were made, while an accelerometer attached to the neck area was used to capture the subglottal waveform. Statistics on our subglottal resonance measurements are given and compared with prior work. Qualitative agreement is shown between the resonator model and diphthong data with time-varying $F2$ for several speakers. Comparison of the second vowel formant and second subglottal formant tracks across all speakers, analysis of the formant spaces spanned by each speaker's vowel data, and a survey of vowel formant data for a sample of the world's languages support the possibility that a speaker's second subglottal resonance divides front and back vowels. Possible implications for theories of vowel inventory structure [e.g., J. Lijencrants and B. Lindblom, Language **48**, 839–862 (1972)] are discussed. [Work supported by NIH Grant DC00075.]

4aSC10. Formant shift in nasalization of vowels. Takayuki Arai (Dept. of Elec. and Electron. Eng., Sophia Univ., Tokyo, Japan and Res. Lab. of Electron., MIT, Cambridge, MA 02139)

We measured the formant shifts of a vowel in the context of a nasal and investigated whether human perception is able to compensate for such shifts. According to the acoustic theory, nasal coupling causes a modification on the spectrum, including formant frequency shift. The first goal of this study is to confirm that the formant frequencies actually shift due to nasalization. Based on several measurements of formant frequencies of various vowels in nasal contexts, we confirmed that the first formant ($F1$) tends to shift in a more central direction when nasalized. In English, vowels should be perceived as the same phoneme regardless of nasalization. In other words, listeners might have the capability to compensate for such formant shifts. The second goal of this study is to examine this compensation effect by a perceptual experiment. For stimuli, we synthesized a nonnasal vowel $V0$ that has the same formant frequencies as a nasalized vowel $V1$. A continuum was also synthesized between $V0$ and the nonnasalized version of $V1$. Results show $V1$ is more correctly identified than $V0$, which suggests the existence of the compensation effect.

4aSC11. Voiceless stop duration under narrow focus and in clear speech. Jeanette A. Ortiz (Dept. of Commun. Sci. and Disord., Northwestern Univ., 2240 Campus Dr., Evanston, IL 60208), Ann R. Bradlow, and Janet B. Pierrehumbert (Northwestern Univ., Evanston, IL 60208)

Talkers typically hyperarticulate when producing individual speech segments under conditions of sentence-level, or narrow focus and of clear speech, which is speech produced in response to known speech perception difficulties on the part of the listener. This study investigated whether narrow focus and clear speech elicit similar changes in the speech signal. A comparison was conducted between the effects of these two phenomena on the durations of voiceless stops /k/ and /t/ in word-initial and word-medial positions of trochaic, disyllabic words. Both narrow focus and clear speech yielded longer closure and aspiration durations when /k/ and /t/ were in word-initial position. For word-medial stops, narrow focus yielded no change in closure duration and a lengthening of aspiration duration for both /k/ and /t/. In contrast, clear speech yielded no change (/t/) or an increase (/k/) in closure duration, and, more interestingly, significant decreases in aspiration duration for both /k/ and /t/. The different effects of narrow focus and clear speech on voiceless stop production suggest that the local hyperarticulation of sentence-level focus and the global hyperarticulation of clear speech probably do not arise from the same underlying organizational structures.

4aSC12. Rosa's roses: Reduced vowels in American English. Edward Flemming (Dept. of Linguist., Stanford Univ., Stanford, CA 94305-2150, flemming@stanford.edu)

Beginning phonetics students are taught that American English has two contrasting reduced vowels, transcribed as [ə] and [ɨ], illustrated by the unstressed vowels in the minimal pair *Rosa's* versus *roses*. However, little seems to be known about the precise nature or distribution of these vowels. This study explores these questions through acoustic analysis of reduced vowels in the speech of 12 American English speakers. The results show that there is a fundamental distinction between the mid central [ə] vowel that can occur in unstressed word-final position (e.g., in *Rosa*), and high reduced vowels that occur in most other unstressed positions, and might be transcribed as [ɨ]. The contrast between pairs like *Rosa's* and *roses* derives from this difference because the word-final [ə] is preserved when an inflectional suffix is added, so the schwa of *Rosa's* is similar to the final vowel of *Rosa*, whereas the unstressed vowel of *roses* is the high [ɨ] reduced vowel quality found elsewhere. So the standard transcription of the reduced vowel contrast is justified, but the widespread use of [ə] to transcribe word-internal reduced vowels is misleading—mid reduced vowels are generally only found in stem-final position.

4aSC13. Child directed speech, speech in noise and hyperarticulated speech in the Pacific Northwest. Richard Wright, Lesley Carmichael, Alicia Beckford Wassink, and Lisa Galvin (Dept. of Linguist., Univ of Washington, Box 354340, Seattle, WA 98195-4340, rawright@u.washington.edu)

Three types of exaggerated speech are thought to be systematic responses to accommodate the needs of the listener: child-directed speech (CDS), hyperspeech, and the Lombard response. CDS (e.g., Kuhl *et al.*, 1997) occurs in interactions with young children and infants. Hyperspeech (Johnson *et al.*, 1993) is a modification in response to listeners difficulties in recovering the intended message. The Lombard response (e.g., Lane *et al.*, 1970) is a compensation for increased noise in the signal. While all three result from adaptations to accommodate the needs of the listener, and therefore should share some features, the triggering conditions are quite different, and therefore should exhibit differences in their phonetic outcomes. While CDS has been the subject of a variety of acoustic studies, it has never been studied in the broader context of the other "exaggerated" speech styles. A large crosslinguistic study was undertaken that compares speech produced under four conditions: spontaneous conversations, CDS aimed at 6–9-month-old infants, hyperarticulated speech, and speech in noise. This talk will present some findings for North American English as spoken in the Pacific Northwest. The measures include $f0$, vowel duration, $F1$ and $F2$ at vowel midpoint, and intensity.

4aSC14. Formants of Japanese function particles. Setsuko Shirai (Dept. of Linguist., Univ. of Washington, P.O. Box 354340, Seattle, WA 98195-4340)

There is much debate about whether vowel centralization results from reduction of undershoot. The results of research conducted by Keating and Huffman (1984) indicated that Japanese vowels in prose were formed closer to the center of vowel chart than in words. However, they did not provide any statistical information and the information about the preceding consonants in the prose. Thus, there was a possibility that the preceding consonants led to the vowel centralization through undershoot. To address this question, I conducted research, in which Japanese function vowels [a, e, o] were compared with content vowels. The results showed that $F1$ of function /a/ following /g/ (average 609.8 Hz) was statistically lower than content /a/ (average 696.3 Hz) [$F(1,65)=73.40$, $p<0.001$]. However, there were no significant differences of $F2$ of /e/ following /d/ and $F2$ of /o/ following /t/ between function and content. At first, it appeared centralization played a role. However, close examination of the results indicated that vowel undershoot was the source of the /a/ centralization. Japanese function /a/ was statistically shorter than content /a/ and there was a significant correlation between duration and normalized vowel displacements for /a/ [Pearson's $r=0.466$, $p<0.001$]. This short duration caused the difference of $F1$ between content /a/ and function /a/.

4aSC15. Word-length and context effects on the acoustics of /ai/. Chandan R. Narayan (Dept. of Linguist., Univ. of Michigan, 4080 Frieze Bldg., Ann Arbor, MI 48109, cnarayan@umich.edu)

The acoustics of American English /ai/ are investigated. This study investigated whether speakers maintain an invariant slope of the $F2$ transition across the vowel shortening effects of increased word length and the vowel lengthening effects of following obstruent voicing. Six speakers recorded minimal /ai/ triplets varying in word length and post-vocalic place and voicing (i.e., hide, Heidi, Heidelberg; hype, hyper, hyperness). Analysis of both temporal and spectral characteristics revealed systematic effects of word length and voicing. With increasing word length /ai/ duration significantly decreased across all variables, except in voiceless contexts where there was not a significant decrease from two- to three-syllable words. Before voiced obstruents $F2$ transition onset increased significantly from one- to two-syllable words with offset frequencies remaining stable. In the three-syllable word condition, the $F2$ transition offset dropped significantly. Importantly, $F2$ transition slope remained stable

across word lengths, which mirrors earlier findings of constant $F2$ slope across different speaking rates (Gay, 1968; Miller *et al.*, 1994). However, $F2$ transition slope was not constant across voicing conditions, being significantly higher before voiceless obstruents than before voiced obstruents. Findings will be compared with earlier work on diphthong acoustics across temporal manipulations.

4aSC16. Co-articulation effects on sound /r/ in cluster words. S. S. Agrawal and Deepak Dhiman (CSIO, CSIR Complex, 2nd Fl., Pusa Campus, New Delhi, India)

In this study the coarticulatory effects on the formant frequencies of sound /r/, one of the most context dependent sounds of Hindi, have been investigated, by taking combinations of the cluster sounds of the forms C_1r, rC_2 and C_3C_4r. In these cluster combinations C_1 is one of the consonants from (/k/,/p/,/g/,/t/,/d/,/m/,/n/,/h/,/b^h/,/w/,/b/,/p^h/), C_2 is one of the consonants from (/y/,/ñ/,/t^h/,/t/,/ʃ/,/m/,/d^h/,/k/,/w/,/n/,/g/,/j/,/b^h/,/tʃ/, /dz/,/ph/,/d/,/s/) and C_3C_4 is one of the combination from (/ʃt/,/st/,/nd/,/n:r/, /nd^h/). A database of 438 cluster words spoken by three male speakers has been prepared. The sound /r/ has been perceptually segmented from these utterances and the first three formants and their corresponding bandwidths have been measured. For C_1r combination, the deviations of the formant frequencies become larger with the order of the formants. The deviations in the formants for rC_2 case are also large and it is nasalized if C_2 is some nasal sound like /ñ/ and /m/. For the third case (C_1C_2r), the formant frequencies do not deviate much as in the earlier two cases but the effect of the frication of the combination /ʃt/ makes it fricative. These results are being used in improving the quality of Hindi Klatt Synthesizer.

4aSC17. Dholuo interdentals: Fricatives or affricates? Evidence from domain-initial strengthening. Rachael Degenshein and Ioana Chitoran (Dartmouth College, Hanover, NH 03755)

The manner of articulation of ⟨th,dh⟩ in Kenya Dholuo is described controversially in the literature. This study uses acoustic data collected from three native speakers of the same dialect to test the hypothesis that ⟨th,dh⟩ exhibit positional variation, in accordance with domain-initial strengthening. The prediction that ⟨th,dh⟩ are realized as affricates in strong position and fricatives in weak position was supported. All three speakers showed longer closure for ⟨th⟩ in pretonic than in non-pretonic position ($F=166.833$, $p<0.01$), and in word-initial than medial position ($F=66.505$, $p<0.01$). In line with the prediction that longer pre-voicing is indicative of a weaker closure, the pre-voicing of ⟨dh⟩ followed the hierarchy: Intonational Phrase (IP) medial word-medial > IP-medial word-initial > IP-initial without focus > IP-initial with focus ($F=116.513$, $p<0.01$). IP-initial ⟨dh⟩ and ⟨d⟩ had similar prevoicing duration (72.120; 72.434) suggesting that the interdental became more stoplike in this position. Furthermore, whereas it is known that vowels are longer before fricatives than before stops, the vowels before ⟨th,dh⟩ were not longer than those before the corresponding stops ⟨t,d⟩ suggesting ⟨th,dh⟩ are stop-like. Using both qualitative and quantitative criteria, Dholuo interdentals are best described as affricates with differences in articulation in prosodically stronger versus weaker positions.

4aSC18. Bimodal schwa: Evidence from acoustic measurements. Noriko Yamane-Tanaka, Bryan Gick, and Sonya Bird (Interdisciplinary Speech Res. Lab., Dept. of Linguist., Univ. of BC, E270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada, nrkyamane@aol.com)

The question of whether schwa is targeted or targetless has been the subject of much debate (Browman *et al.*, 1992; Browman and Goldstein, 1995; Gick, 1999, 2002). Gick (2002) found that there is a pharyngeal constriction during schwa relative to rest position, and concluded that schwa is not targetless. This experiment further showed a ''bimodal'' pattern in schwa in a nonrhotic speaker, indicating that the subject has distinct schwas in lexical words and function words. The present study examines the existence of the ''bimodal'' pattern in schwas in nonrhotic dialects through an acoustic experiment. It is predicted that there is a significant difference in formant values between lexical schwas and function schwas. Results to date indicate a significant difference in them between schwas in lexical versus function words, both between historical schwas and those derived from final /r/ reduction. Data from several additional nonrhotic subjects will be presented. Implications for intrusive r as well as for the phonological treatment of function words will be discussed. [Work funded by NSERC and SSHRC.]

4aSC19. Parametric synthesis of Korean alveolar stops. Gwanhi Yun (Dept. of Linguist., Univ. of Arizona, P.O. Box 210028, Tucson, AZ 85721-0028, ghy@email.arizona.edu)

Korean alveolar stops were synthesized through MITalk system to evaluate the parameters. Because of the insufficiency of reliable acoustic cues to distinguish three-way Korean stops, synthesized stops were expected to lack intelligibility and naturalness. In the first synthesis, AH (amplitude of aspiration during transition from C to the following V), duration of transition, formant frequencies, bandwidths of formants, and $F0$ of the following vowels were employed as main parameters only at two anchor points, i.e., at the burst point and onset of the following vowel. As expected, the plain stops (33%) were not satisfactorily discriminated from tense stops (39%), and naturalness was below chance. However, the discrimination between plain/tense and aspirated stops was not so bad. Further, the intelligibility and naturalness of aspirated stops were so low. Thus, to compensate for lack of reliable parameters, two parameters were readjusted: (1) duration of AH and (2) duration of transition. The identification results showed that the longer AH and transition, the more accurate the perception of aspirated stops. Thus, it indicates that AH and duration of transition might work as a trade-off relation for the perception of aspirated stops, and so a relevant combination of parameters in SynthWork may synthesize more accurate Korean aspirated stops.

4aSC20. One-hand control of a speech synthesizer. Harold A. Cheyne II, Robert E. Beaudoin, Thomas E. von Wiegand, Kenneth N. Stevens (Sensimetrics Corp., 48 Grove St., Ste. 305, Somerville, MA 02144-2500, harold@sens.com), and Patrick M. Zurek (Sensimetrics Corp., Somerville, MA 02144-2500)

The long-term objective of this research is the development of a one-hand-controlled speech synthesizer, to give laryngectomees and other speech-impaired persons a means of producing higher-quality speech with less effort than currently available methods such as an electrolarynx or a text-to-speech system. To demonstrate the feasibility of a one-hand-controlled speech synthesizer, a system was constructed using a hand-held device similar to a pen connected to an articulated arm for measuring six degrees of freedom (three Cartesian and three rotational dimensions) as the user interface to an HLsyn-based speech synthesizer. Through this interface, the user controls parameters for the first three formants, pitch, subglottal pressure, and glottal area. Parameter control was introduced progressively in that order to four participants who underwent training to produce synthesized speech composed of a subset of English phonemes: vowels, semivowels, diphthongs, /h/, and the glottal stop. The complexity of the synthesized speech targets also grew from monosyllabic utterances to short phrases over the training. After training, a separate group of four listeners compared the naturalness and intelligibility of the synthesized speech to the same utterances produced by the participants with a text-to-speech system. [Work supported by NIDCD Grant Number R43 DC006134-01.]

4aSC21. Acceptability of variations in question intonation in natural and synthesized American English. Ann K. Syrdal (AT&T Labs Res., 180 Park Ave., Florham Park, NJ 07932-0971, syrdal@research.att.com) and Matthias Jilka (Univ. of Stuttgart, Stuttgart, Germany)

In a previous study exploring American English question intonation, we found that some speakers deviated considerably from expected question prosody. In this study, we focus on listener-rated acceptability of the various prosodic patterns observed for yes/no and wh questions. A variety of intonational patterns realized in both question utterances recorded from five female and three male professional speakers and in questions synthesized from several TTS voices of both genders was presented to listeners. Subjects judged the acceptability of each utterance in the context of a dialogue between a travel agent and customer. We hypothesized that question utterances with the expected intonational features (phrase-final fall in wh questions, phrase-final rise in yes/no questions) would be rated as more acceptable than question utterances with deviating intonational features, and that this result would hold for both natural and synthetic speech conditions. In addition, following our previous results, we hypothesized that the unexpected intonation pattern of phrase-final falls for yes/no questions would be more acceptable for lower-pitched than for higher-pitched voices. We also varied the prominence of the interrogative pronouns in synthetic wh questions in order to see whether simulating their high intonational prominence in natural wh questions improved the acceptability of synthetic wh questions.

4aSC22. Perceptual consequences of nasal consonant "surrogates" in English: Implications for speech synthesis. Susan R. Hertz (Dept. of Linguist., Cornell Univ., Ithaca, NY 14853 and NovaSpeech, 54 Middlesex Turnpike, Bedford, MA 01730, srh9@cornell.edu), Isaac C. Spence, Thomas F. Church (Cornell Univ., Ithaca, NY 14853), and Richard Goldhor (NovaSpeech, Bedford, MA 01730)

Experiments indicate that non-nasal obstruents in human utterances can be replaced by ''surrogate'' segments, either produced by formant synthesis or recorded from other speakers, with virtually no change in speech quality or speaker identity [Hertz, Proc. IEEE 2002 Workshop on Speech Synthesis (2002)]. While the durational and spectral properties of the surrogate segments must be broadly appropriate to their target context, no speaker-specific tailoring is required. This paper describes follow-on experiments studying the perceptual consequences of replacing nasal consonants in human utterances with surrogate segments from different phonetic contexts, either synthesized or spoken by other speakers. These experiments indicate that the manipulated speech sounds natural when surrogate segment durations, and the formant transitions and nasalization characteristics of adjacent vowels, are appropriate. In certain contexts $F0$ is also perceptually salient. The spectral characteristics of surrogate nasal murmurs are often unimportant. In many cases, the perceived speech quality, phoneme identity, and speaker identity are unaffected even by a surrogate from a phoneme differing from the original. This paper highlights the perceptual results and explains their relevance to hybrid synthesis techniques that employ cross-speaker waveform concatenation and/or integrate waveform concatenation with formant synthesis. Utterances that exemplify these results will be played.

4aSC23. Spectral specification of articulatory parameters. Khalil Iskarous (Haskins Labs., 270 Crown St., New Haven, CT 06511)

It has long been known that there is no unique analytic solution to the inverse problem in speech. However, for any problem with multiple solutions, a unique solution can still be obtained by optimizing some function on the set of possible solutions. Several investigators have done just that for the speech problem. In this work, we return to the analytic approach to ask two questions. First, despite the fact that the spectral frequencies do not uniquely specify the area function, what about articulation do they specify and what additional information is necessary for unique specification? Second, if such information were available, what is the form of the function that specifies the area function given all the necessary spectral information? Based on theorems from network synthesis it is known that the zeros or formant amplitudes can provide the extra information. Assuming that the poles and zeros of the driving point impedance are given (sidestepping the difficulty of obtaining these parameters from a speech signal), a functional relation is presented that uniquely specifies articulatory information given acoustic information. The relation between this method and pole-residue methods will also be discussed. [Work supported by NIH Grant DC-02717.]

4aSC24. Advances in the acoustic correlates for nasals from analysis of MRI data. Tarun Pruthi and Carol Espy-Wilson (Dept. of Elec. and Computer Eng., AVW Bldg., Univ. of Maryland, College Park, MD 20742, tpruthi@glue.umd.edu)

MRI data for nasals is being used to simulate the nasal murmur spectrum in order to improve our understanding of nasals in speech. The data consists of American English nasals /m/, /n/, and /ng/ from one speaker [Story *et al.*, J. Acoust. Soc. Am. **100**, 537–554 (1996)]. A computer simulation model developed in our lab is being used for the purpose [Zhang and Espy-Wilson, J. Acoust. Soc. Am. (2004)]. A sufficiently good match has been obtained between the simulated and real spectra, particularly during the low frequencies (below 2000 Hz). Currently, only the maxillary and sphenoidal sinuses have been included in the model. The match between the two spectra is expected to improve further when frontal and ethomoidal sinuses are included in the model. It is our belief that this study will give us a better understanding of the acoustic manifestations of nasal manner and place in the nasal murmur spectrum and help us achieve our goal of finding speaker-independent acoustic parameters (APs) for them. It would also be interesting to see if this study can help us in finding speaker-dependent APs from the nasal murmur spectrum for our speaker recognition system. [Work supported by NSF Grant No. BCS0236707.]

4aSC25. VTAR: A Matlab-based computer program for vocal tract acoustic modeling. Xinhui Zhou, Zhaoyan Zhang, and Carol Espy-Wilson (Dept. of Elec. and Computer Eng., Inst. of Systems Res., Univ. of Maryland, College Park, MD 20742)

A Matlab-based computer program for vocal tract acoustic response calculation (VTAR) has been developed. Based on a frequency-domain vocal tract model [Z. Zhang and C. Espy-Wilson, J. Acoust. Soc. Am. (2004)], VTAR is able to model various complex sounds such as nasals, rhotics, and liquids. With input in the form of vocal tract cross-sectional area functions, VTAR calculates the vocal tract acoustic response function and the formant frequencies and bandwidths. The user-friendly interface allows directed data input for defined categories: vowels, nasals, nasalized sounds, consonant, laterals, and rhotics. The program also provides an interface for input and modification of arbitrary vocal tract geometry configurations, which is ideal for research applications. [Work supported by NIH Grant 1 R01 DC05250-01.]

4aSC26. Obstruent-sonorant consonant sequences—Analysis by synthesis. Xiaomin Mou (MIT Speech Commun. Group, 77 Massachusetts Ave., 36-513, Cambridge, MA 02139, xmou@mit.edu)

The goal of this work is to develop principles of overlapping gestures in obstruent-sonorant sequences in the word-initial position and sonorant-obstruent sequences in the word-final position. Consonant clusters such as sm in small are phonetically represented as a sequence of individual elements, but the exact perceptual representation is unclear. The modification during the production of these overlapping gestures may be driven partly by perceptual salience and partly by vocal tract aerodynamics. When two consonants occur next to each other, the same gestures may be made as for only one consonant. The aerodynamics of the vocal tract may account for the modification in the timing of the articulators during production and this modification can be incorporated as rules into HLsyn, a higher-level quasiarticulatory speech synthesizer that takes as inputs the pressures and

the flows of the vocal tract. Acoustic information extracted from the speech waveform is mapped into inputs for HLsyn. This analysis by synthesis approach is a method to develop a more precise picture of the planning stage during speech production where the acoustic phonetics must be carefully planned and modified to acheive the correct target sounds. [Work funded by a grant provided by NIH.]

4aSC27. Text-to-phonemic transcription and parsing into mono-syllables of English text. Yugal Jusgir Mullick, S. S. Agrawal, Smita Tayal, and Manisha Goswami (CSIO, CSIRComplex, 2nd Fl., Pusa Campus, New Delhi, India)

The present paper describes a program that converts the English text (entered through the normal computer keyboard) into its phonemic representation and then parses it into mono-syllables. For every letter a set of context based rules is defined in lexical order. A default rule is also defined separately for each letter. Beginning from the first letter of the word the rules are checked and the most appropriate rule is applied on the letter to find its actual orthographic representation. If no matching rule is found, then the default rule is applied. Current rule sets the next position to be analyzed. Proceeding in the same manner orthographic representation for each word can be found. For example, ''reading'' is represented as ''rEdiNX'' by applying the following rules:

r→r move 1 position ahead
ead→Ed move 3 position ahead
i→i move 1 position ahead
ng→NX move 2 position ahead, i.e., end of word.

The phonemic representations obtained from the above procedure are parsed to get mono-syllabic representation for various combinations such as CVC, CVCC, CV, CVCVC, etc. For example, the above phonemic representation will be parsed as rEdiNX-→ /rE/ /diNX/. This study is a part of developing TTS for Indian English.

4aSC28. Signature of prosody in tonal realization: Evidence from Standard Chinese. Yiya Chen (Theoretical and Appl. Linguist. Dept., Edinburgh Univ., Adam Ferguson Bldg., EH8 9LL, UK)

It is by now widely accepted that the articulation of speech is influenced by the prosodic structure into which the utterance is organized. Furthermore, the effect of prosody on $F0$ realization has been shown to be mainly phonological [Beckman and Pierrehumbert (1986); Selkirk and Shen (1990)]. This paper presents data from the $F0$ realizations of lexical tones in Standard Chinese and shows that prosodic factors may influence the articulation of a lexical tone and induce phonetic variations in its surface $F0$ contours, similar to the phonetic effect of prosody on segment articulation [de Jong (1995); Keating and Foureron (1997)]. Data were elicited from four native speakers of Standard Chinese producing all four lexical tones in different tonal contexts and under various focus conditions (i.e., under focus, no focus, and post focus), with three renditions for each condition. The observed $F0$ variations are argued to be best analyzed as resulted from prosodically driven differences in the phonetic implementation of the lexical tonal targets, which in turn is induced by pragmatically driven differences in how distinctive an underlying tonal target should be realized. Implications of this study on the phonetic implementation of phonological tonal targets will also be discussed.

4aSC29. Variation of Taiwanese tones in conversation. Pei-Yu Hsieh and Jane Tsay (Inst. of Linguist., Chung Cheng Univ., 160 San-Xing, Min-Xiong, Chia-Yi 621, Taiwan, ROC, g9015001@ccu.edu.tw)

In laboratory research on tonal coarticulation in Taiwanese, one study [H.-B, Lin, Ph.D. dissertation, Univ. of Connecticut (1988)] reported a perseveratory effect but no anticipatory effect, while another [S.-H. Peng, J. Phonetics **25**, 371–400 (1997)] found a significant anticipatory effect. Peng also found tonal variation due to prosodic positions. Unlike these previous laboratory studies, this study attempts to investigate tonal coarticulation and prosodic effects on Taiwanese tones using natural conversations from the Taiwanese Spoken Corpus (Tsay and Myers, 2004), of which 56 min of recorded conversations were analyzed. Consistent with Lin, the results showed that tone is more affected by the preceding tone than by the following tone. The slope is more influenced by the preceding tone as well. As for prosodic effects, the results confirmed Peng, showing that $F0$ is the lowest in utterance-final position, while in other phrase-final positions it is slightly lower than in non-phrase-final position. This study thus demonstrates the results obtained in the laboratory do indeed carry over into actual conversation.

4aSC30. Contrasting the effects of duration and number of syllables on the perceptual normalization of lexical tones. Valter Ciocca, Alexander L. Francis, and Teresa S.-K. Yau (Div. of Speech & Hearing Sci., Univ. of Hong Kong, 5/F, 34 Hospital Rd., Hong Kong, HKSAR, China, vciocca@hkusua.hku.hk)

In tonal languages, syllabic fundamental frequency ($F0$) patterns (''lexical tones'') convey lexical meaning. Listeners need to relate such pitch patterns to the pitch range of a speaker (''tone normalization'') to accurately identify lexical tones. This study investigated the amount of tonal information required to perform tone normalization. A target CV syllable, perceived as either a high level, a low level, or a mid level Cantonese tone, was preceded by a four-syllable carrier sentence whose $F0$ was shifted (1 semitone), or not shifted. Four conditions were obtained by gating one, two, three, or four syllables from the onset of the target. Presentation rate (normal versus fast) was set such that the duration of the one, two, and three syllable conditions (normal carrier) was equal to that of the two, three, and four syllable conditions (fast carrier). Results suggest that tone normalization is largely accomplished within 250 ms or so prior to target onset, independent of the number of syllables; additional tonal information produces a relatively small increase in tone normalization. Implications for models of lexical tone normalization will be discussed. [Work supported by the RGC of the Hong Kong SAR, Project No. HKU 7193/00H.]

4aSC31. Effects of native language experience on perceptual learning of Cantonese lexical tones. Alexander L. Francis (Dept. of Audiol. and Speech Sci., Purdue Univ., Heavilon Hall, 500 Oval Dr., West Lafayette, IN 47907, francisa@purdue.edu), Valter Ciocca (The Univ. of Hong Kong, Hong Kong, China), and Lian Ma (Beijing Medical Univ., Beijing, China)

In a tonal language syllabic pitch patterns contribute to lexical meaning. Perceptual assimilation models of cross-language perception predict speakers of another tonal language should assimilate Cantonese lexical tones to native tonal categories, affecting identification, discrimination and acquisition. For nontonal language speakers, two possibilities exist. If pitch information is ignored, vowels with different tones should assimilate to the same native category, lowering performance. If tonal information is attended but unused in native categorization, Cantonese tones could be nonassimilable and therefore easily discriminated, and possibly easily identified or learned. Here, native speakers of Mandarin Chinese and American English were trained to identify Cantonese words differing in lexical tone. Discrimination and identification were tested before and after training. Both groups initially performed well on upper register tones (high level, high rising, mid level) and poorly on lower (low falling, low level, low rising). Mandarin listeners improved most at identifying low falling tones; English listeners improved most on low level and low rising tones. Training primarily appeared to improve listeners' ability to make categorical decisions based on direction of pitch change, a feature reportedly under-attended by English speakers, but preferred by Mandarin speakers. [Work supported by research funding from The University of Hong Kong.]

4aSC32. Musical experience and Mandarin tone discrimination and imitation. Terry L. Gottfried, Ann M. Staby, and Christine J. Ziemer (Dept. of Psych., Lawrence Univ., Appleton, WI 54912-0599, Terry.L.Rew-Gottfried@lawrence.edu)

Previous work [T. L. Gottfried and D. Riester, J. Acoust. Soc. Am. **108**, 2604 (2000)] showed that native speakers of American English with musical training performed better than nonmusicians when identifying the four distinctive tones of Mandarin Chinese (high-level, mid-rising, low-dipping, high-falling). Accuracy for both groups was relatively low since listeners were not trained on the phonemic contrasts. Current research compares musicians and nonmusicians on discrimination and imitation of unfamiliar tones. Listeners were presented with two different Mandarin words that had either the same or different tones; listeners indicated whether the *tones* were same or different. Thus, they were required to determine a categorical match (same or different tone), rather than an auditory match. All listeners had significantly more difficulty discriminating between mid-rising and low-dipping tones than with other contrasts. Listeners with more musical training showed significantly greater accuracy in their discrimination. Likewise, musicians' spoken imitations of Mandarin tones (model tokens presented by a native speaker) were rated as significantly more native-like than those of nonmusicians. These findings suggest that musicians may have abilities or training that facilitate their perception and production of Mandarin tones. However, further research is needed to determine whether this advantage transfers to language learning situations.

4aSC33. Teaching tone and intonation with the Prosody Workstation using schematic versus veridical contours. George D. Allen (College of Nursing, Michigan State Univ., East Lansing, MI 48824) and John B. Eulenberg (Michigan State Univ., East Lansing, MI 48824)

Prosodic features of speech (e.g., intonation and rhythm) are often challenging for adults to learn. Most computerized teaching tools, developed to help learners mimic model prosodic patterns, display lines representing the veridical (actual) acoustic fundamental frequency and intensity of the model speech. However, a veridical display may not be optimal for this task. Instead, stereotypical representations (e.g., simplified level or slanting lines) may help by reducing the amount of potentially distracting information. The Prosody Workstation (PW) permits the prosodic contours of both models and users' responses to be displayed using either veridical or stereotypical contours. Users are informed by both visual displays and scores representing the degree of match of their utterance to the model. American English-speaking undergraduates are being studied learning the tone contours and rhythm of Chinese and Hausa utterances ranging in length from two to six syllables. Data include (a) accuracy of mimicking of the models' prosodic contours, measured by the PW; (b) quality of tonal and rhythmic production, judged by native speaker listeners; and (c) learners' perceptions of the ease of the task, measured by a questionnaire at the end of each session.

4aSC34. Tone clarity in mixed pitch/phonation type tones. Jean E. Andruski (Audiol. & Speech-Lang. Pathol., Wayne State Univ., 581 Manoogian Hall, 906 W. Warren Ave., Detroit, MI 48202, j_andruski@wayne.edu)

Lexical tone identity is often determined by a complex of acoustic cues. In Green Mong, a Hmong-Mien language of Southeast Asia, a small subset of tones is characterized by phonation type in addition to pitch height, pitch contour, and duration, which characterize the remaining tones of the language. In tones that incorporate multiple cues to tonal identity, what makes a tone clear, or easy to recognize? This study examines acoustic and perceptual data to address this question. Six native speakers of Green Mong were asked to produce 132 phonological CV words in sentence context, using a conversational speaking style. Seventeen native speakers of the language were then asked to categorize three tones which have similar falling contours, but are differentiated by phonation type (breathy, creaky, and modal). Tokens that were correctly identified by 100% of the listeners were compared with tokens that were relatively poorly identified. Data indicate that the breathy- and creaky-voiced tones are less susceptible to identification errors than the modal-voiced tone. However, the clearest tokens of the three tones are also differentiated by details of pitch contour shape, and by duration. Similarities and differences between acoustic cue values for the best and worst tokens will be discussed.

4aSC35. Temporal and spectral cues in Mandarin tone recognition. Ying-Yee Kong and Fan-Gang Zeng (Dept. of Cognit. Sci. and Dept. of Otolaryngol., Univ. of California–Irvine, Irvine, CA 92697)

Temporal cues contribute significantly to Mandarin tone recognition, but the relevance of formant frequencies is debatable. This study investigates the relative contribution of temporal and formant cues to tone recognition. Three sets of Mandarin stimuli were created. Recorded whispered speech was used to test the contribution of the formant cue, 1- and 8-band noise-modulated speech with 50-Hz envelope cutoff frequency was used to test the contribution of the temporal envelope cue, and 1- and 8-band speech with 500-Hz cutoff frequency was used to test the contribution of the periodicity cue. Four normal-hearing native Mandarin speakers participated. In quiet, subjects achieved an average 87% (8-band) and 82% (1-band) correct with the periodicity cue (500-Hz cutoff), but only 72% (8-band) and 55% (1-band) correct without the periodicity cue (50-Hz cutoff). Whispered speech produced an average 72% correct. From 10 to 10-dB signal-to-noise ratios, this pattern of results was largely preserved, with whispered speech and 8-band conditions without the periodicity cue showing similar performances, which was significantly better than the 1-band without the periodicity cue, but poorer than the 8-band with the periodicity cue. These results suggest all three cues contribute to Mandarin tone recognition, and there is a trade-off among these cues.

4aSC36. Tonal production of early and late Canadian Cantonese bilinguals. Connie K. So (Dept. of Linguist., Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada)

Previous studies show that native speakers have exhibited changes in both phonetic and phonemic characteristics in their L1 production and perception after living in an L2 community for years. Little is known if the phenomenon is also present in the tonal production of native speakers of tone languages who grew up in an English-speaking environment. This research examines both the phonetic and phonemic features of the six Cantonese tones (i.e., vowel duration, pitch contour, and tonal space) produced by two groups of adult Canadian Cantonese bilinguals: early (AOA: less than 1 year) and late (AOA: 10–13 years). The data collected were compared with those of a group of native Cantonese speakers (AOA: greater than 18 years). Analyses revealed that (i) both the early and the late bilinguals produced significantly longer vowel durations than did the comparison group, and that (ii) their pitch contours and tonal space were significantly different from those of the native speakers. In addition, the early bilinguals' tonal productions were more deviant from the native productions than those of the late bilinguals.

4aSC37. Influence of Mandarin tone exposure on the processing of intonation by 14-year-old American adolescents: An fMRI study. Jo-Fu Lotus Lin, Toshiaki Imada, Patricia Kuhl (Dept. of Speech and Hearing Sci. and Inst. for Learning and Brain Sci., Box 357988, Univ. of Washington, Seattle, WA 98195), and Yue Wang (Simon Fraser Univ., Burnaby, BC V5A 1S6, Canada)

This study investigated, for American adolescents, whether the learning of non-native speech contrasts in one prosodic domain (Mandarin Chinese tones) would influence the processing of non-native contrasts in another prosodic domain (Mandarin Chinese intonation). Two groups of 14-year-old American teenagers were tested using the functional magnetic resonance imaging (fMRI) technique, including eight who had received a

two-week Mandarin tone training program and eight with no exposure to Mandarin. Subjects were scanned during identification tasks. Despite their similar behavioral performance on identification of Mandarin intonation, preliminary results showed different cortical activation patterns for the two groups. Teenagers exposed to Mandarin showed similar activation patterns for the Mandarin intonation and Mandarin tone task, with more right-hemisphere activation for intonation, which differed from activation for English intonation. Teenagers without exposure activated similar areas for Mandarin and English intonation. Familiarity with Mandarin tonal contrasts affects brain activation, not only for Mandarin tones but also for Mandarin intonation, suggesting that training effects may transcend levels. [Work supported by NIH (HD 37954) and the Talaris Research Institute.]

4aSC38. Representational specificity of within-category phonetic variation in the mental lexicon. Min Ju and Paul A. Luce (Dept. of Psych., Univ. at Buffalo, Buffalo, NY 14260, mju@buffalo.edu)

This study examines whether within-category variation in voice onset time (VOT) is encoded in long-term memory and affects subsequent word recognition, and whether these effects are modulated by the degree of lexical discriminability. Four long-term repetition-priming experiments were conducted using words containing word-initial voiceless stops varying in VOT. The magnitude of priming was compared between same and different VOT conditions in words with voiced counterparts (pat/bat) and words without voiced counterparts (cow/*gow), and in words with high-frequency counterparts (clue/glue) and words with low-frequency counterparts (cab/gab). If veridical representations of each episode are preserved in memory, variation in VOT should have demonstrable effects on the magnitude of priming. However, if within-category variation is discarded and form-based representations are abstract, variation in VOT should not mediate priming. Further, if lexical discriminability modulates the degree of encoding of within-category variation, words with counterparts and words with high-frequency counterparts should show larger specificity effects than words without counterparts and words with low-frequency counterparts. Implications of these results for the specificity and abstractness of phonetic representations in long-term memory will be discussed.

4aSC39. Predicting the intelligibility of deaf children's speech from acoustic measures. Rosalie M. Uchanski (Dept. of Otolaryngol., Washington Univ. School of Medicine, St. Louis, MO 63110), Ann E. Geers (Univ. of Texas at Dallas, Dallas, TX), Christine M. Brenner (Moog Ctr. for Deaf Education, St. Louis, MO), and Emily A. Tobey (Univ. of Texas at Dallas, Dallas, TX)

A weighted combination of speech-acoustic measures may provide an objective assessment of speech intelligibility in deaf children that could be used to evaluate the benefits of sensory aids and rehabilitation programs. This investigation compared the accuracy of two different approaches, multiple linear regression and a simple neural net. These two methods were applied to identical sets of acoustic measures, including both segmental (e.g., voice-onset times of plosives, spectral moments of fricatives, second formant frequencies of vowels) and suprasegmental measures (e.g., sentence duration, number and frequency of intersentence pauses). These independent variables were obtained from digitized recordings of deaf children's imitations of 11 simple sentences. The dependent measure was the percentage of spoken words from the 36 McGarr Sentences understood by groups of naive listeners. The two predictive methods were trained on speech measures obtained from 123 out of 164 8- and 9-year-old deaf children who used cochlear implants. Then, predictions were obtained using speech measures from the remaining 41 children. Preliminary results indicate that multiple linear regression is a better predictor of intelligibility than the neural net, accounting for 79% as opposed to 65% of the variance in the data. [Work supported by NIH.]

THURSDAY MORNING, 27 MAY 2004 CONFERENCE ROOM E, 8:25 TO 11:55 A.M.

Session 4aSP

Signal Processing Acoustics and Underwater Acoustics: Advances in Sonar and Imaging Techniques Including Interferometric, Synthetic and Tomographic Apertures

John Impagliazzo, Chair
John Impagliazzo Consulting, 2 Kimberly Drive, Wakefield, Rhode Island 02879-3804

Chair's Introduction—8:25

Invited Papers

8:30

4aSP1. Time domain beamforming ASIC for a handheld bistatic imaging sonar system. Alice M. Chiang, Steven R. Broadstone, and John M. Impagliazzo (Teratech Corp., 77-79 Terrace Hall Ave., Burlington, MA 01803)

A high-resolution, handheld imaging sonar system has been developed by Teratech Corporation for the U.S. Navy. This is a 192-channel, dual-frequency bistatic sonar for Navy divers performing search and survey missions for underwater explosives. The goal is to provide the most compact and energy-efficient imaging system for the divers. The low power and small volume are a result of the development of Teratech's Charge Domain Processing (CDP) technology. This technology has led to the development of a low-power 64-channel beamformer chip. As a result, only three beamformer chips are needed for the 192-element array. Until now, implementation of small, low-power sonar systems containing this many elements and forming enough beams to create an image was considered impossible. Test results and acoustic images obtained in Teratech's acoustic test tank will be presented. [Work sponsored by ONR and OSD Small Business Innovative Research Program, Program Manager, Mr. Bruce Johnson, Naval Explosive Ordnance Disposal Technology Division.]

8:50

4aSP2. Interferometry and computer-aided tomography as an acoustic analysis tool. Ron J. Wyber (Midspar Systems Pty. Ltd., 24 Farrer Pl., Oyster Bay, NSW 2225, Australia) and Brian G. Ferguson (Maritime Operations Div. DSTO, Pyrmont, 2009 NSW, Australia)

By using a wide band signal to measure the impulse response of a target it is generally possible to resolve the reflected signal into a sequence of impulses associated with discrete scattering features or structural waves. If two hydrophones, with an appropriate vertical separation, are used to receive the signal from a target, it is possible to determine the vertical position of the source for each impulse in the sequence from the relative phase of the analytic signal received at each of the hydrophones. The amplitude and phase information measured as the target rotates can then be used to form a three-dimensional image in which the position of the strong scattering sources is highlighted. Examples are given showing how the images formed provide information allowing features in the measured impulse response to be related to different scattering mechanisms such as specular target highlights, multiple reflections or structural waves. By characterizing the acoustic hotspots in the image in this way it is possible to provide feedback to the designer of a stealth target quantifying the contribution of individual features on the target to the overall target strength.

Contributed Papers

9:10

4aSP3. Acoustic Mine Imaging (AMI) project: An underwater acoustic camera for use in mine warfare. Colin Ellis and Ed Murphy (Thales Underwater Systems Pty. Ltd., 274 Victoria Rd., Rydalmere, NSW 2116, Australia, colin.ellis@au.thalesgroup.com)

This paper is submitted to detail the advances in sonar and imaging techniques and synthetic apertures being made in Australia by Thales Underwater Systems within a Australian Defence Acquisition Project termed Acoustic Mine Imaging (AMI). This paper will detail the development of the AMI underwater acoustic camera for the detection, classification and characterization of mines and other underwater objects in turbid water where optical imaging is ineffective. It will explain the history of the development from a DSTO concept to the current system. It will detail how the acoustic camera provides real-time milimetric resolution three-dimensional images and how the design has had to meet practical operational constraints to allow seamless mounting onto a small, remotely controlled underwater vehicle for mine disposal operations. An overview of the processing architecture and technical complexity of the AMI will be provided, including detail on the design and development of the required submilimeter transducers and 2D matrix array. The paper will reflect on the technical and operational challenges that had to be addressed. Finally, trial results will be presented that will demonstrate the real-time capability of the acoustic camera in various environments.

9:25

4aSP4. Array processing methods for identifying buried objects. Salah Bourennane (Institut Fresnel/GSM, UMR CNRS 6133, France)

Underwater object identification has been of great interest for a few years to acousticians (detection of boulders) or marines (detection of buried mines). Image and signal processing succeed in identifying objects lying on the sea bottom, however identification of an object buried in sediment remains complex. The purpose of this paper is to develop a complete identification of objects embedded in the sediment using array processing methods. We use a parametric source whose properties are based on the water nonlinear propagation characteristics, the signals are received on a uniform linear antenna. The higher order statistics of the recorded data and the high resolution methods are used to localize the objects and to estimate their ranges. This paper presents also a procedure which estimates discriminant parameters from the signals in order to classify the different objects. The developed algorithms are applied on the experimental underwater acoustic data. The obtained results show the good performance of our algorithms.

9:40

4aSP5. Sonar image enhancement via acoustic color. Nicola Neretti, Nathan Intrator (IBNS, Brown Univ., Providence, RI 02912), and Quyen Huynh (CSS, NSWC, Panama City, FL 32407)

A novel method to enhance synthetic aperture sonar (SAS) images using the acoustic color of the returns from different targets is presented. The method enhances the SAS image using the internal absorption properties of targets in a way that enhances the differences between different types of targets (target signature). First, the algorithm detects the highlights in the sonar image as is found from the beam-formed processed image. These highlights represent edges of targets. Then the analysis is performed only at the echo returns which correspond to edges of targets. For each edge the raw data corresponding to the return from that edge is found. This data which has not yet been pulse-compressed using the pinging signal is analyzed using a detailed frequency representation. Then, the frequency representation of the pinging signal is subtracted from it so that a resulting frequency representation of the difference between the pinging signal and the returning signal is obtained. A normalization of the returned echo is performed before the subtraction to compensate for the lower energy of the returning echo. The same method can be extended to a more detailed analysis of internal structures of targets.

9:55

4aSP6. Bathymetry and seafloor acoustic backscatter imagery with a volume search sonar. Daniel S. Brogan and Christian P. de Moustier (Ctr. for Coastal and Ocean Mapping, Univ. of New Hampshire, 24 Colovos Rd., Durham, NH 03824, daniel.brogan@unh.edu)

Volume search sonars designed for mine hunting applications could be used for environmental sensing, particularly seafloor relief and texture. This capability is explored with a system that transmits a stepped FM pulse over a 243 deg vertical fan beam centered on nadir. It receives with 27 pairs of beams, symmetrically steered about nadir in the fore-aft direction and spaced at 7.16-deg intervals across track. The receive beam pair geometry allows simultaneous views of the seafloor in forward, vertical, and aft profiles. Pulse compression, monopulse processing techniques, and temporal and spatial filtering are used to estimate bathymetry and seafloor acoustic backscatter imagery of a sandy bottom and a muddy bottom. Three monopulse techniques have been investigated: conjugate product, difference over sum, and reduced beamwidth which is the most promising for this application. Results are presented for data collected while surveying at roughly 25 knots, showing the combined effects of acoustic geometry and survey speed on the resolution of the bathymetry and acoustic backscatter imagery and on bottom coverage. [Work supported by NRL Grant No. N00173-00-1-G912.]

10:10–10:25 Break

10:25

4aSP7. Investigations into the application of a new sonar system for assessing fish passage in Alaskan rivers. Deborah Burwen (Alaska Dept. of Fish and Game, 333 Raspberry Rd., Anchorage, AK 99518), Suzanne Maxwell (Alaska Dept. of Fish and Game, Soldotna, AK 99669), and Carl Pfisterer (Alaska Dept. of Fish and Game, Fairbanks, AK 99701)

Over a two-year period (2002–2003) the Alaska Department of Fish and Game (ADFG) has evaluated a relatively new sonar technology at several sites in Alaska to determine its applicability to counting migrating fish in rivers. The new system, called a dual frequency identification sonar (DIDSON), is a high-definition imaging sonar designed and manufactured

by the University of Washingtons Applied Physics Lab for military applications such as diver detection and underwater mine identification. Studies conducted in 2002 in a clearwater stream with high fish passage rates (5000 fish/hour) demonstrated that the DIDSON could provide accurate estimates of fish migrating upstream in rivers. Additional studies conducted in 2002 with live tethered fish indicated that size information could be derived from the cameralike images produced by the DIDSON system at ranges up to 12 m. In 2003, ADFG deployed DIDSON systems to assess fish passage in five Alaskan rivers. In each river, the DIDSON system was deployed in the side-looking configuration typically used in riverine sonar applications for estimating fish passage. Preliminary results from experiments in both 2002 and 2003 indicate that the DIDSON provides significant improvements in our ability to detect, track, and determine the direction of travel of migrating fish in rivers.

10:40

4aSP8. Ultrasound parametric imaging of coral. Emmanuel Bossy (Dept. of Aerosp. and Mech. Eng., Boston Univ., 110 Cummington St., Boston, MA 02215, ebossy@bu.edu), Anne Cohen (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and Robin O. Cleveland (Boston Univ., Boston, MA 02215)

The skeletal density of several massive reef coral species is strongly correlated with ocean temperature, and it has recently been proposed as a paleoclimate proxy offering access to multicentury-long records of sea surface temperature. The presence of annual growth bands (from a few mm up to 1 or 2 cm wide) provides a chronological tool that enables changes in skeletal density to be traced through time. In this work, we investigate the feasibility of imaging the coral skeletal density using ultrasound. As the structure of coral is very close to that of trabecular bone, techniques successfully applied to bone imaging were assessed on coral. Ultrasonic measurements in the MHz frequency range were carried out in the laboratory on machined coral samples, using backscattering and through-transmission configuration. Parametric images were generated from the derivation of ultrasonic parameters such as acoustic impedance, velocity, backscatter coefficient, and attenuation. As x-ray absorption is directly related to the density of the coral, x-ray images were obtained as reference images. Preliminary results show qualitative agreement between x-ray images and ultrasound parametric images, indicating the potential for ultrasound to image coral. [Work supported by NSF through the Centre for Subsurface Sensing and Imaging Systems, Award Number EEC-9986821.]

10:55

4aSP9. Fourier and wavelet domain denoising of active sonar echoes. Peter G. Cable (BBN Technologies, 11 Main St., Mystic, CT 06355-3641), Sheila Shah, and Gary Butler (BBN Technologies, Arlington, VA 22209-3801)

Active sonar classification performance improves significantly when echo-to-background ratios increase above 10–15 dB. To achieve the improved echo waveform fidelity implied by increasing echo-to-background, preclassification processing methods are sought to improve echo waveform estimates. For this purpose a class of nonlinear techniques termed denoising, applied to efficient Hilbert space representations of transient signals, has been shown to yield nearly optimal estimation procedures for noise corrupted signals of unknown smoothness [D. L. Donoho and I. M. Johnstone, Biometrika **81** (1994)]. We have applied several versions of Fourier and wavelet domain denoising to noisy low-frequency target echoes and, for echoes near detection threshold, have demonstrated signal representation improvements equivalent to increases in echo-to-background of 4 dB. The theoretical foundations of denoising, including a new threshold algorithm, will be outlined and measures of performance for waveform estimation will be reviewed and discussed. The experimental methodology used and the results obtained for the test sonar echoes will be summarized and target classification implications of the results obtained from the analysis discussed. [Work supported by ONR.]

11:10

4aSP10. Survey of active sonar simulations. Diana F. McCammon (McCammon Acoust. Consulting, 475 Baseline Rd., RR3, Waterville, NS B0P 1VO, Canada)

In lieu of well-characterized measured data sets, time series simulations are often employed for testing active sonar systems. This paper reviews the simulation techniques and models that are being employed today and discusses some of the issues of fidelity that may arise with regard to range dependence, bistatic geometries, reverberation statistics and bottom interactions.

11:25

4aSP11. Imaging flaws in thin multilayered metal sheets using an array of guided mode transducers. Karl A. Fisher (Lawrence Livermore Natl. Lab., Livermore, CA, 94551)

In a multilayered elastic structure, acoustic energy can be focused into a particular layer by identifying a suitable guided wave mode. Once this guided mode is identified, low profile surface mounted interdigital transducers can be designed to excite preferentially this mode in the multilayered elastic structure. This has led to the development of a novel eight-channel linear array based on guided mode interdigital transducers. The array has been used to generate data for a time domain reconstruction algorithm. Resulting images reveal 1–2-mm length cracks in the center layer of a 300-μm-thick composite steel plate.

11:40

4aSP12. Ultrasonic spacecraft leak location using structure borne noise. Stephen D. Holland, Ron R. Roberts, D. E. Chimenti, and Michael Strei (Ctr. for Nondestructive Eval., Iowa State Univ., 1915 Scholl Rd., Ames, IA 50014)

Leak detection and location in pressure vessels is commonly performed using microphones to detect airborne ultrasound generated by the leak turbulence. For spacecraft, this turbulence is generated outside the spacecraft and cannot be detected inside because the leak velocity is approximately the speed of sound. Instead, to detect and locate leaks in on-orbit spacecraft we monitor leak-generated guided ultrasonic waves within the platelike spacecraft skin. We use cross-correlation to measure the deterministic behavior of the leak-generated noise. Measured leak-into-vacuum noise signals from two adjacent transducers, each correlated with the signal from a third reference transducer, are fed into headphones as a stereo pair. The direction to the leak can be determined by rotating the two transducers, or equivalently by selecting element pairs in a dense array. The leak can be precisely located through triangulation.

Session 4aUW

Underwater Acoustics: Acoustic Propagation and Internal Waves

Peter H. Dahl, Chair
Applied Physics Laboratory, University of Washington, 1013 N.E. 40th Street, Seattle, Washington 98105-6698

Chair's Introduction—7:55

Contributed Papers

8:00

4aUW1. Comparison between ocean-acoustic fluctuations in parabolic-equation simulations and estimates from integral approximations. Stanley M. Flatté and Michael D. Vera[a] (Phys. Dept., Univ. of California at Santa Cruz, Santa Cruz, CA 95064, sflatte@ucsc.edu)

Line-integral approximations to the acoustic path integral have been used to estimate fluctuations due to internal waves. Approximations for the root-mean-square (rms) fluctuation and the bias of travel time, rms fluctuation in vertical arrival angle, and the spreading of the acoustic pulse out to 1000-km range are here compared to estimates from simulations that use the parabolic equation (PE). [See S. M. Flatté and M. D. Vera, J. Acoust. Soc. Am. **114**, 697–706 (2003).] Integral-approximation (IA) estimates of rms travel-time fluctuations were within statistical uncertainty at 1000 km for the Slice89 profile, and in disagreement by between 20% and 60% for the Canonical profile. Bias estimates were accurate for the first few hundred kilometers of propagation, but often disagreed beyond. The PE structure functions of travel time with depth were quadratic for vertical separations of 20 m or less, in qualitative agreement with the IA prediction. Pulse spreads in the PE results were much smaller than predicted by the IA estimates. Implications of these results will be discussed. [Work supported by ONR.] [a] Currently at Scripps Institution of Oceanography.

8:15

4aUW2. Acoustic effects from the horizontal variability spectrum of nonlinear internal waves. Laurel K. Reilly-Raska (Rensselaer Polytech. Inst., Troy, NY 12180, reilll@rpi.edu), James F. Lynch, John A. Colosi (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and William L. Siegmann (Rensselaer Polytech. Inst., Troy, NY 12180)

Recent shallow water experiments suggest that 3-D ocean variability can produce significant intensity variations with a broad spectrum of time scales. In particular, nonlinear internal waves (NIWs) can cause acoustical ducting when the NIW and acoustic propagation directions are nearly at right angles. The sound-speed curvature induced by NIWs play an important role in controlling acoustic focusing, as shown by recent numerical simulations. However, NIWs are not perfect plane waves and an acoustic signal propagating roughly perpendicular to an NIW will encounter a fluctuating duct. We present 3-D numerical simulations in which multiple NIW packets are used to emulate the along-duct variability and its effect on acoustic intensity changes. A rough hierarchy of NIW feature characteristics, based on their intensity influence, will be described. [Work supported by ONR.]

8:30

4aUW3. Blind prediction of broadband coherence time at basin-scales. John L. Spiesberger (Dept. of Earth and Environ. Sci., 240 S. 33rd St., Univ. of Pennsylvania, Philadelphia, PA 19104, johnsr@sas.upenn.edu), Frederick Tappert (Univ. of Miami, Miami, FL 33149), and Andrew R. Jacobson (Princeton Univ., NJ 08544)

A blind comparison with data is made with a model for the coherence time of broadband sound (133 Hz, 17 Hz bandwidth) at 3709 km. Coherence time is limited by changes in the ocean because the acoustic instruments are fixed to the Earth on the bottom of the sea with time bases maintained by atomic clocks. Although the modeled coherence time depends a bit on the difficult problem of correctly modeling relative signal-to-noise ratios, normalized correlation coefficients of the broadband signals for the data (model) are 0.90 (0.83), 0.72 (0.59), and 0.51 (0.36) at lags of 2, 4.1, and 6.2 mins respectively. In all these cases, observed coherence times are a bit longer than modeled. The temporal evolution of the model is based on the linear dispersion relation for internal waves. Acoustic propagation is modeled with the parabolic approximation and the sound speed insensitive operator.

8:45

4aUW4. Analysis of acoustic intensity fluctuations measured after one-way transmission and bistatic target scattering on the New Jersey continental shelf during MAE 2003. Tianrun Chen, Sunwoong Lee, Yisan Lai, Purnima Ratilal, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Acoustic intensity was observed to fluctuate in time and space after both one-way transmission and bistatic scattering from moored targets in the New Jersey continental shelf during the Main Acoustic Clutter Experiment (MAE) 2003 [Lai *et al.*, **114**, 2312 (2003)]. Oceanographic measurements made during the experiment also showed significant temporal and spatial variability in sound speed structure, indicating the presence of oceanic instabilities such as internal waves. Here the theoretical approach of Ratilal and Makris [**114**, 2428 (2003)] is applied to estimate the expected acoustic intensity given the measured statistics of the oceanographic field data. The resulting estimates span a wide range of possible values due to undersampling of the oceanographic data but are still useful in helping to understand the fundamental mechanisms responsible for the observed fluctuations.

9:00

4aUW5. Analytic mean and variance of forward propagated field through random internal waves and subbottom anomalies with Rayleigh–Born scattering. Purnima Ratilal, Tianrun Chen, and Nicholas C. Makris (MIT, 77 Massachuetts Ave., Cambridge, MA 02139)

The mean and variance of the field propagated through a waveguide containing random internal waves or subbottom anomalies is modeled using a modal solution [Ratilal and Makris, J. Acoust. Soc. Am. **114**, 2428 (2003)] that analytically expresses the effects of dispersion, attenuation,

and redistribution of modal energy due to multiple scattering in the forward direction. The scatter function of an elemental volume of inhomogeneity in density and compressibility is modeled using the Rayleigh–Born approximation. Simulations in both continental shelf and deep water waveguides quantitatively show how the relative magnitude of coherent versus incoherent intensity varies with range in typical internal wave fields and subbottom environments.

9:15

4aUW6. Frequency dependence of the sound fluctuations in the presence of internal solitary waves in SWARM'95. Mohsen Badiey (College of Marine Studies, Univ. of Delaware, Newark, DE 19716), Boris Katsnelson, Serguey Pereselkov (Voronezh State Univ., Voronezh 394006, Russia, katz@phys.vsu.ru), and James Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

In this paper the propagation of sound pulses in a shallow-water region in the presence of internal soliton train (IS) traveling approximately across an acoustic track of 15 km is studied. Broadband sound pulses were produced by an airgun (shots with frequency band about 30–120 Hz) and an electromechanical (J15) sound source (LFM sweeps with frequency band 50–200 Hz). Experimental data show significant correlation of sound fluctuations with IS motion. In addition, difference in the excitation source causes different modal energy partitioning in the waveguide. Modeling is carried out within the framework of the theory of vertical modes and horizontal rays (or PE in the horizontal plane). Calculations give a good agreement for temporal fluctuations of the sound energy for the sources used. It is shown that frequency dependence of intensity fluctuations is connected with the frequency dependence of refraction index of horizontal rays due to the 3D nature of these fluctuations. [Work supported by ONR and RFBR, Grant 03-05-64568.]

9:30

4aUW7. Acoustic field propagation through a time-evolving buoyant jet. Steven Finette, Roger Oba (Acoust. Div. Naval Res. Lab., Washington, DC 20375, finette@wave.nrl.navy.mil), Patrick Gallacher, and Steve Piacsek (Naval Res. Lab., Stennis Space Ctr., MS 39529)

Previous theoretical work regarding internal wave/acoustic wave interaction has shown that in the case of soliton (i.e., *nonoverturning*) propagation, significant effects on acoustic propagation occur via both amplitude and phase components. Acoustic flow visualization data obtained recently in the South China Sea [Orr and Mignerey, J. Geophys. Res. **108** (2003)] indicate the presence of internal bores with associated Kelvin–Helmholtz instabilities at the base of the mixed layer. These *overturning* bores are examples of nonhydrostatic dynamics associated with significant vertical acceleration of the fluid. In order to explore the effect of *overturning* on acoustic field propagation, we have used a nonhydrostatic hydrodynamic model to simulate the temporal evolution of internal bores. The initial condition is a stationary front separating two regions characterized by slightly different but homogeneous densities. The front is released at $t=0$ and the gravity-induced flow evolves into an internal bore with associated Kelvin–Helmholtz instabilities, or rotors, behind the leading edge of the bore. Results for transmission loss and signal gain degradation over a frequency range of 200–500 Hz will be presented. [Research sponsored by ONR.]

9:45

4aUW8. A modified approach to scintillation index analysis of intensity. Steven Lutz, David Bradley, and R. Lee Culver (Appl. Res. Lab., Penn State Univ., State College, PA 16804)

Intensity fluctuations of direct and refracted paths measured during a field test near San Clemente Island in August 2002 are presented. Signals used include 20 and 40 kHz CW pulses with 0.14- and 1.0-ms durations. Source to receiver separation was 1 km. Towed CTD data, also collected during the experiment, is presented. Acoustic signals will be analyzed using a matched filter. Temperature variations will be computed as mean-square fluctuations of the index of refraction. The intensity fluctuation records have lower scintillation indices than may be expected from the existing theory [e.g., Uscinsky *et al.*, J. Acoust Soc. Am. **74** (1983)]. However, our frequencies are one to two orders of magnitude higher than those used in experiments for which this theory was developed and our propagation ranges are considerably shorter. More significantly, the acoustic data records are of considerably shorter duration ($\sim$30 s) than previous experiments. To address the issue of the short record's inability to detect larger structures in the water column, filtering is applied to the sound speed field before the calculation of the correlation length (patch size). [Work supported by ONR Code 321US.]

10:00–10:15 Break

10:15

4aUW9. Source localization and inversion in the presence of internal waves. Ralph N. Baer and Michael D. Collins (Naval Res. Lab., Washington, DC 20375, baer@nrl.navy.mil)

It is often possible to localize an acoustic source in a medium with environmental uncertainties by including environmental parameters in the search space [J. Acoust. Soc. Am. **90**, 1410–1422 (1991)]. This approach has recently been applied to problems involving stochastic environmental parameters [J. Acoust. Soc. Am. **114**, 2401 (2003)]. The parameters of the Garrett–Munk spectrum were held fixed, and the source was localized by generating replica fields for various realizations of the internal wave field. In the present work, source localization simulations are performed with the parameters of the Garrett–Munk spectrum included in the parameter space. The possibility of solving an inverse problem to estimate the internal wave parameters is also considered. [Work supported by the ONR.]

10:30

4aUW10. Sea surface effects on reverberation vertical coherence and inverted bottom acoustic parameters in the East China Sea. Ji-Xun Zhou, Xue-Zhen Zhang (School of Mech. Eng., Georgia Inst. of Technol., Atlanta, GA 30332), Peter H. Dahl, and Jeffrey A. Simmen (Univ. of Washington, Seattle, WA 98105)

Wideband reverberation measurements were repeatedly conducted at the center of the ASAEX site in the East China Sea, on 3 and 5 June, 2001. Between these two measurement days, wind speed and rms wave height changed significantly, going from about 1 m/s and 0.1 m to 10 m/s and 0.35 m, respectively. This paper will compare reverberation vertical coherence (RVC) and RVC-inverted equivalent bottom acoustic parameters in frequency range of 100–1200 Hz for the two measurements [J. X. Zhou and X. Z. Zhang, J. Acoust. Soc. Am. **113**, 2204 (2003)]. The difference of equivalent bottom reflection losses, obtained from two measurements, is well explained by the supporting sea surface data. The results show that an inversion of seabottom acoustic parameters from shallow-water long-range reverberation (or sound propagation) should take the surface condition into account, especially for higher sea states and higher frequencies. [Work supported by ONR and NNSF of China.]

10:45

4aUW11. Head waves in the Yellow Sea. Jee Woong Choi and Peter H. Dahl (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

Signals akin to head waves were obtained from explosive sources during the joint U.S. and China Yellow Sea experiment, conducted in August 1996. The sources were deployed at ranges from 0.7 to 37 km in waters approximately 75 m deep. Data were recorded on a 16-element vertical line array with element spacing of 4 m and the deepest element at depth 66 m. This talk will focus on measurements obtained at source ranges less than 1 km, because the amplitude of the head wave arrivals decreases rapidly with range. The arrival time and power spectra were first used to identify and distinguish ground-, water-, and head-wave arrival types. This

latter category was then further studied in the context of pure head waves, noninterfering head waves, and interference head waves. The experimental data are also compared with simulated signals obtained via a Fourier synthesis of a narrow-band complex parabolic equation (PE) field, using the RAM PE code. The overall influence of compressional velocity, velocity gradient, and attenuation in the sediment, and frequency dependency of head wave arrivals are discussed. [Research supported by ONR Ocean Acoustics.]

11:00

4aUW12. Evidence for nonlinear frequency dependence of attenuation in an East China Sea environment. David P. Knobles, Robert A. Koch (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029), James H. Miller, and Gopu R. Potty (Univ. of Rhode Island, Narragansett, RI 02882)

Acoustic data analyzed from two locations in the central East China Sea indicate a nonlinear frequency dependence of the sediment attenuation. The acoustic measurements are in the form of transmission loss (TL) versus range for frequencies in the 50–800-Hz band and time series generated by explosive sources recorded on a vertical line array (VLA). The seabed can be characterized by a sediment that has a significant sand content. It is discovered that a frequency dependence of f^n, with n at least as large as 1.8, is needed to explain both an optimal propagation effect of the TL at the lower frequencies and the temporal spread of the time series measured on the VLA. A linear frequency dependence generates too much loss at the lower frequencies than the reported measurements and impulse responses that have too small of a time spread. Complicating features of the inference of the frequency dependence of the attenuation from forward propagation measurements include the deduction of the source spectrum for individual shots, the source depth, and geoacoustic layering structures that are capable of explaining both measured TL and the time series.

11:15

4aUW13. Low-frequency reverberation signal in shallow water in presence of internal waves. Boris Katsnelson, Sergey Pereselkov (Voronezh Univ., 1, Universitetskaya Sq., Voronezh 394006, Russia, pereselkov@hotmail.com), and Valery Petnikov (General Phys. Inst., Moscow 113000, Russia)

The reverberation of low-frequency (100–400 Hz) sound in shallow water is considered. It is supposed that reverberation signals are formed by summarizing from circle area (radius about 50 km) as result of sound field backscattering on bottom surface. The developed model of reverberation takes into account mode coupling caused by background internal waves in shallow water and depending on direction of backscattered signal. Influence of internal waves on space-time structure of reverberation signal is considered for two cases: single-modal and multi-modal. The computer simulation shows that internal waves lead to significant temporal fluctuations of reverberation signal in both cases. The results of computer simulation are compared with data of experimental research in the Barents Sea. [Work supported by RFBR, Grants 02-02-16509 and 03-05-64568.]

11:30

4aUW14. Environmental uncertainty in modeling of shallow-water reverberation. T. W. Yudichak and D. P. Knobles (Appl. Res. Labs., Univ. of Texas, P.O. Box 8029, Austin, TX 78713-8029)

Simulated monostatic shallow-water reverberation is produced by a broadband model that accounts for propagation and scattering in a waveguide that is horizontally stratified on average, but subject to slightly rough interfaces and small, random spatial fluctuations in density and sound speed. This model has been used to infer parameters characterizing these inhomogeneities (scattering parameters) from experimental data, assuming a specific background geoacoustic profile. In realistic situations, however, the background environment may either not be known with precision or may be too variable to admit a range-independent characterization. It is even possible that several geoacoustically distinct alternative sets of environmental parameters are plausible. This may occur, for instance, when geoacoustic inversion is used to determine a background profile. In this presentation, the effects of environmental uncertainties, including the existence of alternative geoacoustic profiles, on scattering parameters inferred from reverberation data are examined. [Work supported by ONR.]

11:45

4aUW15. Numerical investigation of rough surface scattering in a fetch-limited environment using a parabolic equation model. Richard Ead (Autonomous Undersea Vehicles Div. Naval Undersea Warfare Ctr., Newport, RI 02841, kbsmith@nps.navy.mil), Kevin B. Smith (Naval Postgrad. School, Monterey, CA 93943), and Mohsen Badiey (Univ. of Delaware, Newark, DE 19716)

When compared to the open ocean, the littoral environment is responsible for a greater dispersion of acoustic energy per volume due to scattering from the sea surface, sea floor, topographic features, and disharmonious regions within the volume. Of particular interest is the modeling of acoustic energy as it is reflected and dispersed by a rough sea surface based on an empirical fetch-limited ocean wave spectrum, and how such modeling compares with measured data. Following the method of Tappert and Nghiem-Phu [J. Acoust. Soc. Am. Suppl. 1 **77**(2), S101 (1985)], a formulation for scattering from a rough surface has been developed that includes results from analysis of data [Hasselmann *et al.*, JPO **10**, 1264–1280 (1980)] collected during the Joint North Sea Wave Observation Project. The formulation is integrated into the Monterey–Miami Parabolic Equation Model [Smith, J. Comput. Acoust. **9**, 243–285 (2001)]. The model is validated by comparing to a unique set of experimental data collected in a fetch-limited region. Arrival time and arrival-angle fluctuations are examined and compared with field data.

Session 4pAA

Architectural Acoustics: Alternative Acoustic Environments for Performing Arts Presentations

Christopher Jaffe, Cochair
Jaffe Holden Acoustics, Inc., 114 A Washington Street, Norwalk, Connecticut 06854

Neil Rolnik, Cochair
Jaffe Holden Acoustics, Inc., 114 A Washington Street, Norwalk, Connecticut 06854

Chair's Introduction—2:45

Invited Papers

2:50

4pAA1. Engaging spaces: Intimate electro-acoustic display in alternative performance venues. Curtis Bahn and Stephan Moore (Arts Dept./iEAR Studios, Rensselaer Polytechnic Inst., 110 8th St., Troy, New York, NY 12180, crb@rpi.edu)

In past presentations to the ASA, we have described the design and construction of four generations of unique spherical speakers (multichannel, outward-radiating geodesic speaker arrays) and Sensor-Speaker-Arrays, (SenSAs: combinations of various sensor devices with outward-radiating multichannel speaker arrays). This presentation will detail the ways in which arrays of these speakers have been employed in alternative performance venues—providing presence and intimacy in the performance of electro-acoustic chamber music and sound installation, while engaging natural and unique acoustical qualities of various locations. We will present documentation of the use of multichannel sonic diffusion arrays in small clubs, ''black-box'' theaters, planetariums, and art galleries.

3:10

4pAA2. Alternative acoustic environments for the generation of reverberation. Alexander U. Case (fermata audio + acoustics, P.O. Box 1161, Portsmouth, NH 03802)

The musicians and engineers who create popular recorded music view reverberation as a signal processing effect to be added to any and all elements of a multitrack production. Devices such as digital reverbs, spring reverbs, and plate reverbs are tools of the recording trade, synthesizing reverblike sounds for performance through loudspeakers. Acoustic reverberation makes its way into recorded music through the use of a reverb chamber. A small room is used to generate reverb. With cubic volume well below that of a performance hall, it works the ''other side'' of the Sabine equation, being built of highly sound reflective materials. A purpose-built room for the generation of reverb is a luxury not many studios can afford. Clever use of stairwells, bathrooms, and basements is easier on the recording studios balance sheet. This work evaluates the repurposing of these alternative spaces for the generation of reverb in popular recorded music.

3:30

4pAA3. Not your grandfather's concert hall. Russell Cooper (Jaffe Holden Acoustics, Inc., 114A Washington St., Norwalk, CT 06854, rcooper@jhacoustics.com), Richard Malenka (Carnegie Hall, New York, NY 10019-3210), Charles Griffith (Polshek Partnership, New York, NY 10014), and Steven Friedlander (Auerbach, Pollack, Friedlander, New York, NY 10018-5509)

The opening of Judy and Arthur Zankel Hall on 12 September 2003, restores Andrew Carnegie's original 1891 concept of having three outstanding auditoriums of different sizes under one roof, and creates a 21st-century venue for music performance and education. With concerts ranging from early music to avant-garde multimedia productions, from jazz to world music, and from solo recitals to chamber music, Zankel Hall expands the breadth and depth of Carnegie Hall's offerings. It allows for the integration of programming across three halls with minifestivals tailored both to the size and strengths of each hall and to the artists and music to be performed. The new flexible space also provides Carnegie Hall with an education center equipped with advanced communications technology. This paper discusses the unique program planned for this facility and how the architects, theatre consultants, and acousticians developed a design that fulfilled the client's expectations and coordinated the construction of the facility under the floor of the main Isaac Stern Auditorium without having to cancel a single performance.

3:50

4pAA4. A composer's eye for room acoustics. Alvin Lucier (Dept. of Music, Wesleyan Univ., 42 Pinewood Terrace, Middletown, CT 06457, alucier@wesleyan.edu)

The physical acoustician has a number of techniques that can be used to document the acoustic signature of an enclosed space. A composer may approach this matter from a different perspective. This paper discusses and sonically illustrates several physical acoustic aural approaches to developing the signature of a room using unamplified speech and music as source signals. Other methods involving multiple echograms and natural impulse sources are also described. Portions of a score composed for five instruments will be played. In this piece, parts were written for four acoustic instruments as well as for the room in which the performance was presented.

4:10

4pAA5. Open secrets. Malcolm Holzman (Hardy Holzman Pfieffer & Assoc., 902 Broadway, New York, NY 10010, mholzman@hhpa.com), Eve Beglarian (EVBVD Music, New York, NY 10011), and Curtis Bahn (Rensselaer Polytechnic Inst., Troy, NY 12180-3590)

In 1998, the Schools of Architecture and Humanities at the Rensselaer Polytechnic Institute initiated a collaborative project between students of both schools to develop a musical performance piece that would be presented in an environment to be designed and built in the Art Gallery of the School of Architecture. Students from both Schools were to work together on developing the content and conceptualizing the built environment. Architects Malcolm Holzman and Nestor Bottino of the firm of HHPA were brought in as advisors to the project, as was Composer Eve Beglarian of Twisted Tutu. Curtis Bahn, Director of the iEAR Studio at RPI, was Academic Coordinator for the project, together with Beth Weinstein, Professor, School of Architecture, and Grethe Holby, who directed the performance. This paper will review the academic *raison d'etre* for organizing such a program, discuss the history of the project as it developed, and present a video tape of portions of the finished performance.

Contributed Paper

4:30

4pAA6. Urban spaces adopted by the performing arts. Alexander U. Case (fermata audio + acoustics, P.O. Box 1161, Portsmouth, NH 03802)

No celebration of acoustics in New York City would be complete without mention of the countless alternative spaces used for the performance of music, dance, and other art forms. New York's countless train stations, parks, alleys, and overpasses provide popular venues for artists to make sound. Lincoln Center and Carnegie get all the glory, but it's the streets and train stations that do all the work. This work surveys some of these alternative performance spaces and their acoustic contribution to the performance.

THURSDAY AFTERNOON, 27 MAY 2004 LIBERTY 3, 1:00 TO 5:30 P.M.

Session 4pAB

Animal Bioacoustics: Infrasonic Communication by Animals: Signal Propagation, Generation, Reception and Function

Edmund Gerstein, Cochair
Leviathan Legacy, Inc., 1318 S.W. 14th Street, Boca Raton, Florida 33486

Roger S. Payne, Cochair
Ocean Alliance, 191 Weston Road, Lincoln, Massachusetts 01773

Chair's Introduction—1:00

Invited Papers

1:05

4pAB1. Infrasound propagation and coupling between air, soil, and water. Henry Bass, Carrick Talmadge, Craig Hickey (Natl. Ctr. for Physical Acoust., Univ. of Mississippi, Univ., MS 38677, pabass@olemiss.edu), and Francine Desharnais (DRDC Atlantic, Dartmouth, NS, Canada)

Propagation of low-frequency sound is enhanced by very low absorption in air, water, and soil. Some examples of absorption expected in air, water, and soil are presented. Reduction in sound level between source and receiver is typically the result of spreading. In all media of interest, layering can give rise to trapped waves that decay as cylindrical waves as opposed to spherical waves. Examples of the effect of layering on propagation in the atmosphere will be given for day/night conditions. At very low frequency, scattering by rough terrain, moderate size turbulence, and waves have a reduced effect on signal amplitude and phase. Turbulence carried by the wind can result in pressure variations at the receiver that can sound like acoustic signals. An effective receiving system must deal with that wind noise. Identifying the mechanism for communications at low frequency is complicated by the exchange of acoustic energy between air, soil, and water. Some characteristics of this coupling and losses that might be encountered along different paths will be discussed.

1:25

4pAB2. Eavesdropping on elephants. Katy Payne (Bioacoust. Res. Prog., Cornell Lab of Ornithology, 159 Sapsucker Woods Rd., Ithaca, NY 14850, kp17@cornell.edu)

The Elephant Listening Project is creating an acoustic monitoring program for African forest elephants, an endangered species that lives in dense forests where visual censusing is impossible. In 2002, a $2\frac{1}{2}$-month continuous recording was made on an array of autonomous recording units (ARUs) surrounding a forest clearing in the Central African Republic. Each day between 10 and 160 forest elephants (*Loxodonta cyclotis*), the subjects of Andrea Turkalo's 13-year demographic study, were present on the clearing.

Thousands of vocalizations were recorded, most of which contained infrasonic energy. The calls were located in space using software developed in the Bioacoustics Research Program. During daylight hours simultaneous video recordings were made. GPS time-synchronization of video cameras and the ARUs made it possible to identify the elephants responsible for many calls and to examine associated circumstances and behaviors. Recordings were also made on a second acoustic array, permitting a preliminary estimate of propagation and an indication of source level for selected elephant calls. Automatic detection of elephant calls is increasing the feasibility of analyzing long acoustic recordings, and paving the way for finer-tuned analyses, with an ultimate goal of describing forest elephants' acoustic repertoire.

1:45

4pAB3. Elephant low-frequency vocalizations propagate in the ground and seismic playbacks of these vocalizations are detectable by wild African elephants (*Loxodonta africana*). Caitlin E. O'Connell-Rodwell, Jason D. Wood, Roland Gunther, Simon Klemperer (Dept. of Geophys., Stanford Univ., Stanford, CA 94305, ceoconnell@stanford.edu), Timothy C. Rodwell, Sunil Puria, Robert Sapolsky (Stanford Univ., Stanford, CA 94305), Colleen Kinzley (Oakland Zoo), Byron T. Arnason (Tezar, Inc., Austin, TX 76235), and Lynette A. Hart (Univ. of California, Davis, CA 95616)

Seismic correlates of low-frequency vocalizations in African and Asian elephants propagate in the ground at different velocities, with the potential of traveling farther than their airborne counterparts. A semblance technique applied to linear moveouts on narrow-bandpass-filtered data, coupled with forward modeling, demonstrates that the complex waves observed are the interference of an air wave and a Rayleigh wave traveling at the appropriate velocities. The Rayleigh wave appears to be generated at or close to the elephant, either by coupling through the elephant's body or through the air near the body to the ground. Low-frequency elephant vocalizations were reproduced seismically and played back to both a captive elephant and to elephant breeding herds in the wild, monitoring the elephants' behavioral responses, spacing between herd members and time spent at the water hole as an index of heightened vigilance. Breeding herds detected and responded appropriately to seismically transmitted elephant warning calls. The captive studies promise to elucidate a vibrotactile threshold of sensitivity for the elephant foot. Elephants may benefit from the exploitation of seismic cues as an additional communication modality, thus expanding their signaling repertoire and extending their range of potential communication and eavesdropping beyond that possible with airborne sound.

2:05

4pAB4. The role of infrasounds in maintaining whale herds. Roger S. Payne (Ocean Alliance, 191 Weston Rd., Lincoln, MA 01773)

For whales and dolphins a basic social unit is the herd. In several species, herds have been observed to maintain the same speed, direction, and membership overnight, and while swimming in waters of near-zero visibility—evidence that individuals can stay together using nonvisual cues. The most likely such cue is sound. If whale herds are held together with sound, yet we define herds as groups of whales seen moving together, then we are using visual criteria to judge what is an acoustic phenomenon, and our conclusions about a most basic unit of cetacean social structure, the herd, are at least incomplete, and, quite possibly, worthless. By calling herds, heards, we remind ourselves that sound controls herd size. We then consider that some whale infrasound can propagate across deep water at useful intensities (even in today's ship-noise-polluted ocean) for thousands of kilometers. The distance to which blue and fin whale sounds propagate before falling below background noise is given, and the possible advantages these whales obtain from such sounds is explored. The conclusion is that by sharing information on food finds infrasonically, fin and blue whales may have developed a way to divide up the food resources of an entire ocean.

2:25

4pAB5. Baleen whale infrasonic sounds: Natural variability and function. Christopher W. Clark (Lab. of Ornithology, Cornell Univ., 159 Sapsucker Woods Rd., Ithaca, NY 14850, cwc2@cornell.edu)

Blue and fin whales (*Balaenoptera musculus and B. physalus*) produce very intense, long, patterned sequences of infrasonic sounds. The acoustic characteristics of these sounds suggest strong selection for signals optimized for very long-range propagation in the deep ocean as first hypothesized by Payne and Webb in 1971. This hypothesis has been partially validated by very long-range detections using hydrophone arrays in deep water. Humpback songs recorded in deep water contain units in the 20–l00 Hz range, and these relatively simple song components are detectable out to many hundreds of miles. The mid-winter peak in the occurrence of 20-Hz fin whale sounds led Watkins to hypothesize a reproductive function similar to humpback (*Megaptera novaeangliae*) song, and by default this function has been extended to blue whale songs. More recent evidence shows that blue and fin whales produce infrasonic calls in high latitudes during the feeding season, and that singing is associated with areas of high productivity where females congregate to feed. Acoustic sampling over broad spatial and temporal scales for baleen species is revealing higher geographic and seasonal variability in the low-frequency vocal behaviors than previously reported, suggesting that present explanations for baleen whale sounds are too simplistic.

2:45

4pAB6. Do manatees utilize infrasonic communication or detection? Edmund Gerstein (Florida Atlantic Univ., Charles E. Schmidt College of Sci., 777 Glades Rd., Boca Raton, FL 33431 and Leviathan Legacy, Inc., Boca Raton, FL 33486), Laura Gerstein (Leviathan Legacy, Inc., Boca Raton, FL 33486), Steve Forsythe (Naval Undersea Warfare Ctr., Newport, RI 02841), and Joseph Blue (Leviathan Legacy, Inc., FL 32806)

Some researchers speculate Sirenians might utilize infrasonic communication like their distant elephant cousins; however, audiogram measurements and calibrated manatee vocalizations do not support this contention. A comprehensive series of hearing tests conducted with West Indian manatees yielded the first and most definitive audiogram for any Sirenian. The manatee hearing tests were also the first controlled underwater infrasonic psychometric tests with any marine mammal. Auditory thresholds were measured from 0.4 to 46 kHz, but detection thresholds of possible vibrotactile origin were measured as low as 0.015 kHz. Manatees have short hairs

on their bodies that may be sensitive vibrotactile receptors capable of detecting particle displacement in the near field. To detect these signals the manatee rotated on axis, exposing the densest portion of hairs toward the projector. Manatees inhabit shallow water where particle motion detection may be more useful near the water's surface, where sound pressures are low due to the Lloyd mirror effect. With respect to intraspecific communication, no infrasonic spectra have been identified in hundreds of calibrated calls. Low source levels and propagation limits in shallow-water habitats suggest low-frequency manatee calls have limited utility over long distances and infrasonic communication is not an attribute shared with elephants.

3:05–3:20 Break

3:20

4pAB7. Low-frequency sounds and amphibious communication in *Hippopotamus amphibious*. William E. Barklow (Framingham State College, Framingham, MA 01760)

Hippos make sounds in both air and underwater, and, with their heads in an amphibious position (eyes and nostrils above water but mouth and throat below), are able to transmit sounds to both media simultaneously. Hippos on the surface respond to the surface component by calling. Hippos underwater consistently surface and call in a chorus that can spread in air from one territory to the next for many kilometers. They produce several low-frequency, high-amplitude (100 dB re: 20 μPa) sounds. The grunt, their most common call, has a 30- to 60-Hz fundamental, and the huff and some tonal sounds end with an abrupt drop in frequency to 20- to 30-Hz. These sounds are usually given amphibiously, but the high-pass filter characteristics of shallow water attenuates the low frequencies of the underwater component. Hippos also emit these and other sounds when they are completely submerged. These are inaudible in air, but they produce a fountain on the surface accompanied by a 10- to 20-Hz sound. They also produce this effect with plosive blows underwater without other sounds. Similar ''bubble blasts'' have been reported in gray whales. The function of these sounds is not clear, but they may facilitate long-distance ''chain chorusing.''

3:40

4pAB8. Ground sounds: Seismic detection in the golden mole. Peter M. Narins (Dept. of Physiol. Sci., UCLA, 621 Charles E. Young Dr. S., Los Angeles, CA 90095-1606, pnarins@ucla.edu) and Edwin R. Lewis (Univ. of California, Berkeley, CA 94720)

The Namib Desert golden mole is a nocturnal, surface-foraging mammal, possessing a massively hypertrophied malleus which presumably confers low-frequency, substrate-vibration sensitivity through inertial bone conduction. Foraging trails are punctuated with characteristic sand disturbances in which the animal's head dips under the sand. The function of this behavior is not known but it is thought that it may be used to obtain a seismic fix on the next mound to be visited. To test this, we measured the local seismic vibrations both on the top of a mound and on the flats. The spectrum recorded on the flats shows a relatively low-amplitude peak at about 120 Hz, whereas the spectral peak recorded from the mound is nearly 17 dB greater in amplitude and centered at 310 Hz. This suggests that mounds act as seismic beacons for the golden moles that would be detectable from distances corresponding to typical intermound distances of 20–25 m. In addition, out of the 117 species for which data are available, these golden moles have the greatest ossicular mass relative to body size (Mason, personal communication). Functionally, they appear to be low-frequency specialists, and it is likely that golden moles hear through substrate conduction. [Work supported by NIH.]

Contributed Papers

4:00

4pAB9. Interactive patterns of vocal communication in African elephant herds (*Loxodonta africana*). Caitlin E. O'Connell-Rodwell (Dept. of Geophys., Bio-X Program, Stanford Univ., Clark Bldg. E-150, 318 Campus Dr., Stanford, CA 94305-5437, ceoconnell@stanford.edu), Megan T. Wyman, Lynette A. Hart, and Shay Redfield (Univ. of California, Davis, CA 95616)

This study examines the interactive nature of vocalizations produced within African elephant herds during waterhole visits at Etosha National Park, Namibia. The temporal organization and physical characteristics of 1025 vocalizations, documented within 14.8 h of acoustic field recordings, were analyzed using a variety of statistical tests. Temporal distribution analyses indicated that calls were clumped in time (Kologorov–Smirnov, average significant $p = 0.001\,58$) and the majority of the calls occurred after herds began the process of departing a waterhole (Wilcoxon signed rank, $p = 0.0012$). In addition, 84% of all calls began within 30 s of another call. Based on the analysis of the temporal patterning of calls, a conversational series of vocalizations was defined as calls separated by less than 30 s of silence. The majority of calls in a conversational series are overlapping or contiguous, indicating that multiple elephants are involved in the exchange. A variety of physical characteristics (duration, frequency, frequency modulation, and frequency at peak amplitude) was compared across three different call types. These conversational call patterns may function to maintain intragroup cohesion and coordination as well as intergroup resource sharing and competition reduction. [Work supported by: Namibia Nature Foundation, Etosha Ecological Institute, UC Davis Block, Jastro Shields, and Faculty Research Grants.]

4:15

4pAB10. How many rumbles are there? Acoustic variation and individual identity in the rumble vocalizations of African elephants (*Loxodonta africana*). Joseph M. Soltis, Anne Savage (Animal Programs, Disney's Animal Kingdom, P.O. Box 10000, Lake Buena Vista, FL 32830, Joseph.Soltis@disney.com), and Kirsten M. Leong (Cornell Univ., Ithaca, NY)

The most commonly occurring elephant vocalization is the rumble, a frequency-modulated call with infrasonic components. Upwards of ten distinct rumble subtypes have been proposed, but little quantitative work on the acoustic properties of rumbles has been conducted. Rumble vocalizations ($N = 269$) from six females housed at Disney's Animal Kingdom were analyzed. Vocalizations were recorded from microphones in collars around subject necks, and rumbles were digitized and measured using SIGNAL software. Sixteen acoustic variables were measured for each call, extracting both source and filter features. Multidimensional scaling analysis indicates that there are no acoustically distinct rumble subtypes, but that there is quantitative variation across rumbles. Discriminant function analysis showed that the acoustic characteristics of rumbles differ across females. A classification success rate of 65% was achieved when assigning unselected rumbles to one of the six females (test set = 64 calls) according to the functions derived from the originally selected calls (training set = 205 calls). The rumble is best viewed as a single call type with graded variation, but information regarding individual identity is encoded in female rumbles.

4:30

4pAB11. Anatomy of infrasonic communication in baleen whales: Divergent mechanisms of sound generation in mysticetes and odontocetes. Joy S. Reidenberg and Jeffrey T. Laitman (Ctr. for Anatomy and Functional Morphology, Mail Box 1007, Mount Sinai School of Medicine, New York, NY 10029-6574, joy.reidenberg@mssm.edu)

Cetaceans produce sounds at opposite ends of the frequency spectrum. The laryngeal role in odontocete sound production (echolocation, communication) remains unclear. Mysticete infrasonics are presumed to be laryngeal in origin, but production mechanisms are unknown. To address this, we examined postmortem larynges in 6 mysticete species (3 genera) and compared them to our odontocete collection (20 species/15 genera). Results indicate that the rostral portion of the odontocete larynx is elongated, narrow, rigid, and normally positioned intranarially. This portion of the mysticete larynx is comparatively shortened, open, pliable, and in Megaptera may be retracted from its intranarial position. Internally, mysticete vocal folds are thick, paired, and oriented horizontally, compared with the thin, usually unpaired, and vertically oriented odontocete fold. Mysticetes may generate low frequency sounds via pneumatically driven fold vibrations, which then pass to attached laryngeal sac walls, through overlying throat pleats, to water. Rorqual mysticetes may also vibrate paired corniculate flaps while regulating airflow into the nasal region. Infrasonic pulses may pass through adjacent soft palate, skull, or nasal cartilages to water. Laryngeal anatomy in mysticetes and odontocetes appears highly divergent. These morphological differences may correlate to adaptations for producing infrasonic (mysticete) or ultrasonic (odontocete) communication. [Work supported by ONR:N00014-96-1-0764, ONR:N00014-99-1-0815, and AMNHSOF.]

4:45

4pAB12. Spectrogram analysis of low to mid frequency marine mammal clicks. George E. Ioup, Juliette W. Ioup, James P. Larue (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148), Natalia A. Sidorovskaia (Univ. of Louisiana at Lafayette, Lafayette, LA 70504-4210), Stan A. Kuczaj, Grayson H. Rayborn, and Christopher D. Walker (Univ. of Southern Mississippi, Hattiesburg, MS 39406)

Previous investigators have proposed explanations for some sperm whale click structure and pointed out that the separation of individual pulses within the click might be used to determine approximately the size of the sperm whales. Recently, Mohl *et al.* [J. Acoust. Soc. Am. **114**, 1124–1154 (2003)] have shown that echo-location click structure is highly dependent on the received angle. In data measured by the Littoral Acoustic Demonstration Center using bottom-moored hydrophones in the northern Gulf of Mexico in the summers of 2001 and 2002, rich click structures were observed in the spectrograms of many click trains, some of which exhibit strikingly consistent spectral nulls across the train. Although this structure in the spectra could be due to propagation effects, investigations to date suggest this possibility is highly unlikely, as discussed in the next abstract. Therefore it is at least plausible that the structure could be used to identify individual animals. This is known to be a difficult problem in the case of sperm whales because of the angle dependence of at least some of their clicks. These difficulties are discussed, as is the possible use of the spectrograms of the clicks to identify individuals. [Research supported by ONR.]

5:00

4pAB13. Propagation modeling for sperm whale acoustic clicks in the northern Gulf of Mexico. Natalia A. Sidorovskaia, Ilya A. Udovydchenkov, Irina I. Rypina (Dept. of Phys., Univ. of Louisiana at Lafayette, Lafayette, LA 70504-4210), George E. Ioup, Juliette W. Ioup (Univ. of New Orleans, New Orleans, LA 70148), Jerald W. Caruthers (Univ. of Southern Mississippi, Stennis Space Ctr., MS 39529), Joal Newcomb, and Robert Fisher (Naval Res. Lab., Stennis Space Ctr., MS 39529-5004)

Simulations of acoustic broadband (500–6000 Hz) pulse propagation in the northern Gulf of Mexico, based on environmental data collected as a part of the Littoral Acoustic Demonstration Center (LADC) experiments in the summers of 2001 and 2002, are presented. The results of the modeling support the hypothesis that consistent spectrogram interference patterns observed in the LADC marine mammal phonation data cannot be explained by the propagation effects for temporal analysis windows corresponding to the duration of an animal click, and may be due to a uniqueness of an individual animal phonation apparatus. The utilization of simulation data for the development of an animal tracking algorithm based on the acoustic recordings of a single bottom-moored hydrophone is discussed. The identification of the bottom and surface reflected clicks from the same animal is attempted. The critical ranges for listening to a deep-water forging animal by a surface receiving system are estimated. [Research supported by ONR.]

5:15

4pAB14. Sperm whale identification using self-organizing maps. Juliette W. Ioup and George E. Ioup (Dept. of Phys., Univ. of New Orleans, New Orleans, LA 70148)

Self-organizing maps (SOMs) are a neural network technique for clustering data with similar features. Sperm whale phonations are among the many sounds that can be heard in the Littoral Acoustic Demonstration Center underwater acoustic data from three bottom-mounted hydrophones in the northern Gulf of Mexico during the summer of 2001. When more than one whale is present, it would be useful to associate particular clicks or click trains with specific whales. SOMs are employed using various features including the time series itself, Fourier transform coefficients, and wavelet transform coefficients. Preliminary results with a relatively small data set will be presented. [Research supported by ONR.]

Session 4pAO

Acoustical Oceanography and Animal Bioacoustics: D. Van Holliday Special Session on Acoustical Measurements of Marine Organisms II

Joseph D. Warren, Chair
Natural Science Division, Southampton College, 239 Montauk Highway, Southampton, New York 11968

Contributed Papers

1:00

4pAO1. Multifrequency acoustic observations of zooplankton in Knight Inlet, B.C. Mark V. Trevorrow (Defence R&D Canada—Atlantic, P.O. Box 1012, Dartmouth, NS B2Y 3Z7, Canada, Mark.Trevorrow@drdc-rddc.gc.ca), David L. Mackas (Inst. of Ocean Sci., Sidney, BC, Canada), and Mark C. Benfield (Louisiana State Univ., Baton Rouge, LA 70803)

A collaborative investigation of midwater zooplankton aggregations in a coastal fjord was conducted in November 2002. Midwater aggregations of zooplankton in a coastal fjord were sampled and mapped using a calibrated, three-frequency (38, 120, and 200 kHz) vessel-based echo-sounder system, a multinet towed zooplankton net (BIONESS), and a high-resolution *in situ* camera system (ZOOVIS). Dense daytime layers of euphausiids and amphipods near 70- to 90-m depth were found in the lower reaches of the inlet, especially concentrated by tidal flows around a sill which rises above the layer. Quantitative euphausiid and amphipod backscattering measurements, combined with *in situ* species, size, and abundance estimates, were found to agree closely with recent size- and orientation-averaged fluid–cylinder scattering models produced by Stanton *et al.* Also, *in situ* scattering measurements of physonect siphonophores were found to have a much stronger low-frequency (38 kHz) scattering strength, in agreement with a simple bubble scattering model. [Work supported by Dr. J. Eckman, ONR code 322BC.]

1:15

4pAO2. Abundance and diel migration of demersal plankton on a tropical coral reef: An acoustic pilot study. Peter B. Ortner, S. Jack Stamates, Shailer R. Cummings (NOAA/AOML, 4301 Rickenbacker Cswy., Miami, FL 33149), Sharon L. Smith, and Peter V. Lane (UM/RSMAS/MBF, Miami, FL 33149)

Coral reefs serve as the habitat for demersal mesozooplankton and small fishes that migrate into overlying waters at night but spend daylight hours within the reef in part because that habitat provides protection against visual predators. These movements structure energy, mass, and nutrient exchange between the reef habitat and the surrounding waters. Information to date has, however, been predominately qualitative and has not taken advantage of recent advances in biological oceanographic sampling instrumentation. Moreover, what sampling has been done has not been rigorously coupled to synoptic time series of ambient oceanographic conditions. To begin to fill this critical gap we deployed for two 1-month-periods multi-frequency acoustic (TAPS) and optical integrated environmental sensor packages to continuously measure and record the abundance and size distribution of organisms while concomitantly measuring water column chlorophyll, fluorescence, transmittance, temperature, and salinity. During the two periods of high-resolution sampling, and in the intervening six months, we also deployed bottom-mounted ADCP units yielding both vertical current structure and backscatter amplitude distributions. Both video footage and traditional net or pump samples were obtained for ground-truth purposes. [Support provided by the National Marine Fisheries Service/Southeast Fisheries Science Center.]

1:30

4pAO3. Near-shore krill distribution around Livingston Island, Antarctica. Joseph D. Warren (Southampton College, 239 Montauk Hwy., Southampton, NY 11968) and David A. Demer (Southwest Fisheries Sci. Ctr., La Jolla, CA 92037)

The near-shore regions of Livingston Island, Antarctica were studied during the austral summers of 2002 and 2004 to measure the Antarctic krill population and factors which control their distribution. These waters serve as the main feeding grounds for the seasonally resident fur seal and penguin populations of Cape Shirreff. Shallow and highly variable bathymetry in this area precludes surveying from large ships, therefore a specially modified, 19-ft inflatable vessel was used to conduct meteorological and acoustic backscatter surveys of the near-shore region. Concurrently, a 300-ft vessel conducted an overlapping, off-shore survey and collected acoustic backscatter, hydrographic, and net tow data. Results from this study indicate that the near-shore waters contained higher concentrations of acoustic scatterers (believed to be krill) and that the distribution of these animals was related to the presence of two deep submarine canyons that flank Cape Shirreff. Hydrographic data indicate that the near-surface waters near these canyons contained elevated levels of temperature and oxygen. During the 2004 study, several buoys equipped with echosounders were deployed at these canyons to better study their effect on the krill population.

1:45

4pAO4. Seasonal variation in the distribution of *Calanus finmarchicus* and its predators observed through multifrequency acoustics in the Irminger Sea. Sophie Fielding, Alexander T. Mustard (Southampton Oceanogr. Ctr., European Way, Southampton SO14 3ZH, UK), Cairistiona I. H. Anderson, Andrew S. Brierley (Univ. of St. Andrews, St. Andrews, Fife KY16 8LB, UK), and Michael R. Heath (FRS Marine Lab., Torry, Aberdeen AB11 9DB, UK)

The seasonal, and in particular, winter distribution of *Calanus finmarchicus* and its predators were investigated using multifrequency (38–600 kHz) acoustic backscatter measurements in the Irminger Sea, North Atlantic, as part of Marine Productivity, a UK contribution to GLOBEC. The distribution of Calanus, which over-winters deeper than 500 m, well below the acoustic penetration depth of a frequency suitable for detecting targets of approximately 0.2 cm from the surface, was examined using a 600-kHz RDI Acoustic Doppler Current Profiler (ADCP) mounted on a lowered CTD frame. A net system, for species information, and an optical plankton counter (OPC), for numerical abundance, was deployed at each station to ground-truth the 600-kHz acoustic backscatter data. Model-predicted profiles of acoustic backscatter, calculated using OPC derived Calanus abundances and a fluid-filled cylinder acoustic scattering model, were consistent with the observed 600-kHz profiles. Calanus predators, euphausiids, e.g., Meganyctiphanes norvegica and fish, e.g., Myctophids, were observed using a surface towed Simrad EK500 multifrequency echosounder. Seasonal variations in the depth of sound scattering layers at each frequency were observed within the Irminger Sea, which suggests that the Calanus predators may be deepening their winter depth to coincide with the maximum over-wintering Calanus abundances.

2:00

4pAO5. *In situ* and experimental observations of the relationships between euphausiid orientation, vessel lights, and acoustical scattering. Mark C. Benfield (Oceanogr. and Coastal Sci., Coastal Fisheries Inst., Louisiana State Univ., Baton Rouge, LA 70803), Michelle L. Ashton (Louisiana State Univ., Baton Rouge, LA 70803), Mark V. Trevorrow (Defence Res. and Development Canada Atlantic, Dartmouth, NS B2Y 3Z7, Canada), and David L. Mackas (Inst. of Ocean Sci., Sidney, BC V8L 4B2, Canada)

During a collaborative investigation of zooplankton aggregations near a coastal fjord (Knight Inlet, British Columbia), surveys were conducted using down-looking echosounders and a digital imaging system (ZOOVIS). Two surveys conducted at night revealed that the presence or absence of external illumination on the vessel had a pronounced influence on measured acoustical scattering. The extremely short latency between shifts in illumination and changes in acoustical scattering suggested that differences in the orientations of scatterers were responsible for this phenomenon. High-resolution, *in situ* images of euphausiids (*Euphausia pacifica*) from ZOOVIS indicated that these organisms were present in a range of orientations ranging from horizontal to vertical, relative to the incident acoustical beam. Theoretical scattering models based on digitizations of *in situ* ZOOVIS images suggest that the magnitude of the observed changes in scattering, in response to altered illumination, may be accounted for by changes in euphausiid orientation. [Research supported by the ONR, Code 322BC.]

2:15

4pAO6. Behavioral observations of *in situ* copepods with a multibeam sonar. Jules Jaffe (Marine Physical Lab, Scripps Inst. of Oceanogr., La Jolla, CA 92093-0238, jules@mpl.ucsd.edu), Amatzia Genin, and Moti Ohevia (Hebrew Univ., Eilat 88103, Israel)

In order to count and track copepods in three dimensions a multibeam sonar has been developed which operates at a frequency of 1.6 MHz. The sonar uses an analog version of a crossed array in order to localize animals in 2–4-liter volumes. The system uses an 8×8 beam configuration to localize animals at subcentimeter accuracy in range and centimeter accuracy across track. Sensitivity was measured in a test tank via the injection of animals, and it was found that the system was capable of measuring reflections of animals as small as 1–2 mm at ranges of at least 2 m. In order to measure the behavior of the animals under different environmental conditions *in situ*, experiments were conducted in the Gulf of Eilat. Here, the sonar was mounted on a tripod and aimed horizontally. Processed results yielded over 200 000 individual tracks that lasted from 2 to 40 s. The results demonstrated that zooplankton retain their depth under up- and down-welling currents of 1–2 cm/s (as measured with injected fluorescein dye). This is unlike their response in the horizontal, where the animals are passively swept with the currents.

2:30–2:45 Break

2:45

4pAO7. Multifrequency analyses of fish distributions in the northwest Atlantic. J. Michael Jech (NEFSC, 166 Water St., Woods Hole, MA 02543, michael.jech@noaa.gov)

Routine acoustical surveys for estimating Atlantic herring (*Clupea harengus*) population abundance have been conducted on Georges Bank during the autumn spawning season from 1998 to present. Acoustical data are collected with a Simrad EK500 scientific echo sounder operating at 12 or 18, 38, and 120 kHz, and split-beam (the 12-kHz system is a single beam) transducers. Biological measurements and verification of acoustical scatterers are obtained with a pelagic trawl. Acoustical data are evaluated (scrutinized) manually to remove noise, faulty bottom detections, and to classify acoustical backscattering to species. Species classification is currently subjective, and is based on the experience of the scientists and trawl catches. Objective species classification and automated fish density and abundance estimates are an obvious goal for fisheries surveys using advanced technologies. Classification methods using relationships among frequency-dependent volume backscattering strengths, such as presence–absence and combination–permutation, are described and presented. Results indicate that while classification using these methods and acoustical information alone is not robust, these methods highlight backscattering patterns within aggregations and have the potential to characterize backscattering patterns observed in fisheries acoustics data. [Work supported by NOAA Fisheries and ONR.]

3:00

4pAO8. The acoustic environment of the Florida manatee: Correlation with level of habitat use. Jennifer L. Miksis-Olds (Univ. of Rhode Island, Grad. School of Oceanogr., Narragansett, RI 02882, jmiksis@gso.uri.edu), James H. Miller (Univ. of Rhode Island, Narragansett, RI 02882), and Peter L. Tyack (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The Florida manatee is regularly exposed to high volumes of vessel traffic and other human-related noise pollutants because of their coastal distribution. Quantifying specific aspects of the manatees' acoustic environment will allow for a better understanding of how these animals are responding to both natural and human induced changes in their environment. Acoustic recordings and transmission loss measurements were made in two critical manatee habitats: seagrass beds and dredged basins. Twenty-four sampling sites were chosen based on the frequency of manatee presence in specific areas from 2000–2003. Recordings were composed of both ambient noise levels and transient noise sources. The Monterey-Miami Parabolic Equation Model (MMPE) was used to relate environmental parameters to transmission loss, and model outputs were verified by field tests at all sites. Preliminary results indicate that high-use grassbeds have higher levels of transmission loss compared to low-use sites. Additionally, high-use grassbeds have lower ambient noise in the early morning and later afternoon hours compared to low-use grassbeds. The application of noise measurements and model results can now be used to predict received levels, signal-to-noise ratios, and reliable detection of biologically relevant signals in manatee habitats and in the many different environments that marine mammals live.

3:15

4pAO9. Killer whale caller localization using a hydrophone array in an oceanarium pool. Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com), Charles F. Greenlaw, Duncan E. McGehee, and D. Van Holliday (BAE Systems, San Diego, CA 92123)

A system to localize calling killer whales was designed around a ten-hydrophone array in a pool at SeaWorld San Diego. The array consisted of nine ITC 8212 and one ITC 6050H hydrophones mounted in recessed 30×30 cm^2 niches. Eight of the hydrophones were connected to a Compaq Armada E500 laptop computer through a National Instruments DAQ 6024E PCMCIA A/D data acquisition card and a BNC-2120 signal conditioner. The system was calibrated with a 139-dB, 4.5-kHz pinger. Acoustic data were collected during four 48–72 h recording sessions, simultaneously with video recorded from a four-camera array. Calling whales were localized by one of two methods, (1) at the hydrophone reporting the highest sound exposure level and (2) using custom-designed 3-D localization software based on time-of-arrival (ORCA). Complex reverberations in the niches and pool made locations based on time of arrival difficult to collect. Based on preliminary analysis of data from four sessions (400+ calls/session), the hydrophone reporting the highest level reliably attributed callers 51%–100% of the time. This represents a substantial improvement over attribution rates of 5%–15% obtained with single hydrophone recordings. [Funding provided by Hubbs-SeaWorld Research Institute and the Hubbs Society.]

3:30

4pAO10. Contribution of active and passive acoustics to study oceanographic processes feeding whales in a critical habitat of the St. Lawrence Estuary. Yvan Simard, Nathalie Roy (Maurice Lamontagne Inst., Fisheries and Oceans Canada, 850 route de la Mer, Mont-Joli, QC G5H 3Z4, Canada), Yvan Simard, and Cédric Cotté (Univ. du Québec a Rimouski, Rimouski, QC G5L 3A1, Canada, Yvan_Simard@uqar.qc.ca)

The head of the main channel of the continent in eastern Canada is the site of particular oceanographic processes that are responsible for the creation of a persistent feeding ground regularly visited by baleen whales from the Atlantic for centuries. Multifrequency acoustics coupled with ADCP and hydrographic measurements has been used to map the krill and capelin aggregations in 3D and visualize their local concentration process under tidal forcing and upwelling at the channel head. The krill scattering layers, pumped into the area by the strong two–layer estuarine circulation, appear to be concentrated during flood by tidal currents forced against the slopes and upwelling, to which depth-keeping krill is reacting by swimming down. Capelin also tends to concentrate on slopes and neighboring shallows. This highly recurrent process generates rich patches that are contributing with the mean circulation to make this area the richest krill aggregation in Northwest Atlantic. This critical habitat is located in a major continental seaway. Passive acoustics techniques are explored to locate whale calls and map the use of this area in continuing months, especially by blue and fin whales, with the aim of understanding their movements to improve their protection.

3:45

4pAO11. Sonar off-axis target classification by an echolocating dolphin. Patrick Moore (Space and Naval Warfare Systems Ctr. San Diego, 53560 Hull St. Code 2351, San Diego, CA 92152-5001), Lois Dankiewicz (SAIC BioSolutions, San Diego, CA 92110), and Dorian Houser (BIOMIMETICA, La Mesa, CA 91942)

Dolphin echolocation has evolved over millions of years under selection pressures imposed by a selective niche. The complexity and effectiveness of dolphin echolocation for detection and classification of objects within that niche has useful application to U. S. Naval objectives. In these environments, Navy dolphins are likely to first encounter targets on the edge of their sonar beam during a search. It is unknown, however, if target classification is possible from the off-axis (OA) information alone, or whether a more centrally focused interrogation is necessary. This talk addresses the initial findings of an animal detecting two different targets (cylinder and sphere) presented OA (left and right). Data collection methods will be presented. Outgoing echolocation clicks and echoes are digitized and stored to a PC for acoustic characterization using a high-speed Integrated Circuits Systems, Ltd. 32 channel A/D card, sampling 24 calibrated monitor hydrophones and analog filter-amplifiers arranged in a hemispherical support web in front of the animal. Emitted signals analyzed for various acoustic characteristics are discussed as well as detection performance. Since this is an on-going study, available results to date will be presented.

4:00

4pAO12. Unusual spatiotemporal patterns in fish chorusing. Gerald L. D'Spain and Lewis P. Berger (Marine Phys. Lab., Scripps Inst. of Oceanogr., La Jolla, CA 93940-0704)

An unusual fish chorusing behavior has been observed in nighttime underwater acoustic recordings during the Summer off the Southern California coast. Some characteristics of these choruses, i.e., increases in the ocean sound levels by factors of 2 to 5 for 10 to 20 s followed by 15- to 20-s periods of lower levels, repeating every 30 to 40 s all throughout the night, have been described previously. Recently, in reanalyzing data collected during a set of experiments in which Van Holliday participated, we discovered that the choruses along one 25-km stretch of coastline have characteristics analogous to "The Mexican Wave" performed by spectators at sporting events worldwide. Each cycle of the chorus begins in waters off the Mexican coast and the region of chorusing propagates upcoast until the fish just south of the mouth of San Diego harbor reach a chorus peak 16 to 20 s after their Mexican counterparts downcoast. This pattern repeats at 30- to 40-s intervals. The speed of this upcoast migration is 100 times faster than the 12 m/s human waves in stadia, approaching the 1.5 km/s speed of sound in water. [Work supported by ONR.]

4:15

4pAO13. Marine Mammal Active Sonar Test 2004 (MAST 2004). Peter J. Stein, Amy Vandiver (Sci. Solutions, Inc., 99 Perimeter Rd., Nashua, NH 03063-1325, pstein@scisol.com), Geoffrey S. Edelson (BAE Systems, Nashua, NH 03061-0868), Adam S. Frankel (Marine Acoust., Inc., Annapolis, MD 21401), and Christopher W. Clark (Cornell Univ., Ithaca, NY 14850)

The Marine Mammal Active Sonar Test was conducted in January 2004 off the central California coast during the gray whale migration. The purpose of the test was to collect data to significantly advance active sonar detection, classification, and tracking of marine mammals out to ranges of 1 mile. The R/V NEW HORIZON was moored in the migration path. Two different sonar systems operating between 210 and 220 dB were deployed off the vessel, the IMAPS phased array sonar system that operates from 20–30 kHz and the MAST mechanical system that uses rotating parabolic transducers operating from 30 to 40 kHz. Marine mammal observers were deployed on the bluffs overlooking the experiment and aboard the NEW HORIZON. The observers tracked the whales using electronic theodelites, providing ground truth for the sonar systems. They also looked for any severe reactions from the animals, and called for a shutdown if any marine mammals were observed within 100 m of the sonar. Here, we discuss the experiment and the results to date, including any reaction of the whales to the sonar system. We will also touch on the legal battle to conduct the experiment, and lessons learned.

4:30

4pAO14. Sources of uncertainty in Doppler sonar measurements of fish speed. Cristina D. S. Tollefsen and Len Zedel (Dept. of Phys. and Phys. Ocean., Memorial Univ. of NF, St. John's, NF A1B 3X7, Canada)

A 250-kHz, 30-kHz bandwidth coherent Doppler sonar was evaluated to determine sources of uncertainty in fish speed measurements. Three separate tests were undertaken: (1) towtank tests using styrofoam balls to simulate fish, (2) tank tests with live free-swimming fish, and (3) field tests with wild free-swimming fish. The standard deviation in a single speed estimate was 9 cm s^{-1} for styrofoam balls, 10–11 cm s^{-1} for swimming fish observed from a dorsal aspect, and 19 cm s^{-1} for swimming fish observed from a caudal aspect. The variation in precision was primarily due to the different signal-to-noise ratio (SNR) in each test: a larger SNR resulted in a smaller standard deviation. Doppler speed estimates were compared with independent estimates of target speed where possible. An accuracy of ± 4 cm s^{-1} was typical of Doppler speed estimates in all the experiments.

Session 4pBB

Biomedical Ultrasound/Bioresponse to Vibration: Contrast Agents, Cavitation and Lithotripsy

Charly Thomas, Chair

Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

1:00

4pBB1. Inertial cavitation dose produced *ex vivo* in rabbit ear arteries with optison. Juan Tu, Andrew Brayman, and Thomas Matula (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St. Seattle, WA 98105, matula@apl.washington.edu)

Ultrasound-induced inertial cavitation (IC) effects were studied *ex vivo* in rabbit ear arteries with the addition of ultrasound contrast agents (UCAs). Ears were removed from New Zealand white rabbits immediately after being euthanized under a protocol approved by the University of Washington IACUC. The auricular arteries were perfused with varying concentration of UCA (Optison) in saline and exposed to 1.155-MHz pulsed high-intensity focused ultrasound (HIFU) with constant PRF (10 Hz), pulse length (20 cycles), and total treatment time (20 s). Experiments were performed for variable peak negative acoustic pressure ($P-$) (from 0.19 to 3.31 Mpa) and Optison volume concentration (0% [saline only], 0.1%, 0.2%, 0.5%, and 1%). Cavitation activity was quantified by IC Dose (cumulated root-mean-squared [rms] broadband noise amplitude in a particular band in the frequency domain). The results showed that (1) IC activity was induced much more easily with the addition of Optison, even at low volume concentration, such as 0.1%. (2) IC dose increased significantly with the increasing acoustic pressure and Optison concentration. (3) Higher concentrations of Optison decreased the IC threshold. [Work supported by NIH 8RO1 EB00350-2.]

1:15

4pBB2. Signal-to-noise ratio and attenuation of Optison® microbubbles in blood as a function of imaging frequency. Paolo Zanetti, Constantin-C. Coussios, and Ronald A. Roy (Boston Univ., 110 Cummington St., Boston, MA 02215, pzanetti@bu.edu)

Using an active cavitation detector (ACD), the power backscattered by various concentrations of Optison® microbubbles (signal) was compared to the power backscattered by a 50% hematocrit suspension of red blood cells in saline (noise), as a function of imaging frequency (5–30 MHz). A theoretical model, based on direct experimental measurements of the size distribution of Optison® microbubbles, was developed to predict the signal-to-noise ratio (SNR) of microbubbles in blood, assuming no interactions between the populations of scatterers. The SNR was shown experimentally to decrease with increasing imaging frequency up to a point where Optison® no longer provided image enhancement. Measurements of the SNR were repeated in a suspension of 0.8% hematocrit, which has the same backscattering coefficient as a 50% hematocrit suspension. The SNR for Optison® in 50% hematocrit was found to be lower than for the 0.8% hematocrit suspension at all frequencies, suggesting that the large number density and close proximity of the red blood cells inhibits the acoustic response of the microbubbles. Measurements of sound attenuation through suspensions of red blood cells with or without Optison® were also obtained, indicating that the microbubbles barely contribute to the overall attenuation. [Work supported by the NSF and the ASA.]

1:30

4pBB3. *In vitro* characterization of echogenic liposomes by acoustic scattering at 3.5–15.0 MHz. Tyrone M. Porter, Christy K. Holland, Saurabh Datta (Dept. of Biomed. Eng., Univ. of Cincinnati, 6166 MSB, 231 Albert Sabin Way, Cincinnati, OH 45267-0586, tyrone.porter@uc.edu), Shaoling Hwang, Robert C. MacDonald (Northwestern Univ., Evanston, IL 60208), and David D. McPherson (Northwestern Univ., Chicago IL 60611)

Echogenic liposomes (ELIP) are phospholipid vesicles that are being developed as ultrasound contrast agents and as vehicles for drug delivery. These particles, when conjugated with antibodies or peptides, can be used for targeted diagnostic imaging and therapy. Evaluating the acoustic properties of ELIP will allow for optimization of its utility as an ultrasound contrast agent. The diameter of the liposomes ranges from 0.25–10.0 μm. In order to measure the backscattering coefficient *in vitro*, the backscattered power was compared to that reflected from a perfect reflector, a planar air–water interface. The backscatter coefficient and attenuation of ELIP were evaluated as a function of concentration ($0.8–9.9\times10^8$/ml) and frequency (3.5–15.0 MHz). By comparing the values of measured backscatter coefficient to a theoretical model that treats the gas within the liposomes as a free air bubble, a size estimate of the encapsulated gas is provided. Finally, calculation of the scattering to attenuation ratio (STAR) gives a value with which to evaluate the efficacy of ELIP as a contrast agent in cardiovascular ultrasound imaging. The backscatter coefficient shows promise as a sensitive method for determining whether the liposomes are left intact or if they are destroyed during imaging.

1:45

4pBB4. Bubble translation and deformation induced by ultrasound radiation force. Yurii A. Ilinskii, G. Douglas Meegan, Evgenia A. Zabolotskaya (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029, zhenia@arlut.utexas.edu), and Stanislav Y. Emelianov (Univ. of Texas, Austin, TX 78712-1084)

Measurement of small-bubble dynamics has been proposed for the remote evaluation of tissue elasticity [Erpelding *et al.*, Proc. IEEE Ultrasonics Symp., 554–557, 2003]. For example, a microbubble can be produced within the cornea during femtosecond laser surgery and its response to a pulsed ultrasonic radiation force can be measured. The bubble's translation, deformation, and oscillation can be directly related to the mechanical properties of surrounding tissue information that is required for optimization of the surgical procedure. In the work reported here, a model was developed to predict the translation and deformation of an initially spherical bubble in a soft viscoelastic medium as induced by radiation pressure. The extent of bubble translation and deformation is dictated by the elastic stress and viscous forces that oppose the radiation pressure. Numerical simulations predict static, periodic, and transient translation of the bubble in response to continuous, periodic, and pulsed waveforms, respectively. The model also predicts the deformation of the bubble. The results indicate increased deformations with increased bubble translations. Overall, the model can be used to determine the local shear modulus and viscosity of the medium based on measurements of gas-bubble displacement. [Work supported by ARL:UT IR&D.]

2:00

4pBB5. Linear contrast agent detection through low frequency manipulation of high frequency scattering properties. Rune Hansen and Bjorn A. Angelsen (Dept. of Circulation and Medical Imaging, Norwegian Univ. of Sci. and Technol., Norway)

In medical ultrasound imaging, contrast agents in the form of encapsulated gas bubbles are injected into the blood to enhance the scattered blood signal which is weak compared to the scattered tissue signal. Obtaining blood information is, from a medical diagnostic point of view, very helpful. Due to the strong linearly back-scattered tissue signal, contrast imaging today relies on the nonlinear scattering properties of the added gas bubbles. These harmonic techniques typically have important limitations in sensitivity, specificity, and image range resolution. The present paper proposes a new method applying the total scattered contrast signal for image reconstruction, thus largely overcoming the problems encountered in contrast harmonic ultrasound imaging techniques. In the new method, contrast signals and tissue signals are differentiated applying a simple pulse subtraction technique which cancels or significantly reduces the scattered tissue signal, whereas the scattered contrast signal, due to assisting transmitted low frequency pulses altering the acoustic scattering properties of the contrast agent, is preserved in this process. The main mechanism through which this imaging technique selects the contrast agent signal is the linear resonant properties of the contrast bubble and the new method is thus mainly a linear contrast agent detection technique.

2:15

4pBB6. Interaction between therapeutic ultrasound propagation and cavitation bubbles. Marko Liebler, Thomas Dreyer, and Rainer Riedlinger (Institut fuer Hoechstfrequenztechnik und Elektronik/Akustik, Univ. of Karlsruhe, Germany, marko.liebler@ihe.uka.de)

In medical applications of high intense focused ultrasound using pressure pulses or continuous wave signals, cavitation is considered to play a significant role for physical and biological effects. To further develop therapeutic applications it is essential to improve the understanding of these cavitation related effects. In this paper a numerical model is presented to simulate the interactions between ultrasonic waves and cavitation bubbles. The FDTD model is based on a two-phase continuum approach for bubbly liquids and combines nonlinear ultrasound propagation with cavitation bubble activity. Experimental and numerical investigations are presented demonstrating the influence of cavitation bubbles on ultrasound propagation. Measurements with a fiber optic hydrofone for pulsed piezoelectric transducers show significant variations in focal pressure waveforms after the first tensile phase of the wave for different gas content. It is supposed that these changes are caused by cavitation effects. Calculations with different bubble densities confirm these experimental results and demonstrate that the first positive pressure part of the wave is not affected by bubble activity. Increasing the gas content leads to a truncated tensile part followed by augmented pressure oscillations. Further on, simulation results for the evolution and impact of cavitation bubble clouds in CW applications are presented.

2:30

4pBB7. Design of multi-frequency cavitation fields for spatial control of cavitation. Sham D. Sokka, Thomas P. Gauthier, and Kullervo Hynynen (Dept. of Radiol., Brigham and Women's Hospital, Harvard Med. School, Boston, MA 02115)

Cavitation has been implicated as the primary mechanism for a whole host of emerging applications. In all these applications, the main concern is to induce cavitation in perfectly controlled locations in the field; this means specifically to be able to achieve cavitation threshold at the geometrical focus of the transducer without stimulating its near field. In this study, we develop multi-frequency methods to preferentially lower the cavitation threshold at the focus relative to the rest of the field. Three families of multi-frequency driving waveforms are evaluated in a bubble model incorporating rectified diffusion. The results from the optimal waveform analysis are verified by experiment. Finally, the performance of the rest of the acoustic field in suppressing cavitation when cavitation is induced at the focus is investigated theoretically and checked experimentally. This study shows that multi-frequency phased arrays could be used to precisely control cavitation. Cavitation threshold is proved to be almost 1.5 times higher in the near field than at the focus. The concept of cavitation field is introduced and complements cavitation studies concentrating on the focal behavior only.

2:45

4pBB8. Simulation of an acoustically excited bubble near a simulated "cell." Sheryl M. Gracewski (Dept. of Mech. Eng., Univ. of Rochester, Rochester, NY 14620, grace@me.rochester.edu), Hongyu Miao, Diane Dalecki (Univ. of Rochester, Rochester, NY 14620), and Morton W. Miller (Univ. of Rochester, Rochester, NY 14642)

A variety of independent studies have reported increased bioeffects such as hemolysis and hemorrhage induced by high-intensity ultrasound when ultrasound contrast agents are present. Therefore, to better understand the role of cavitation, one-, two-, and three-dimensional models have been developed to investigate the interactions between ultrasonically excited bubbles and model ''cells.'' First, a simple one-dimensional model based on the Rayleigh–Plesset equation was used to estimate upper bounds for strain, strain rate, and areal expansion of a simulated red blood cell. Then, two- and three-dimensional boundary element models were developed (with DynaFlow Inc.) to obtain simulations of asymmetric bubble dynamics in the presence of rigid and deformable spheres. A spherical ''cell'' near an ultrasonically excited bubble was modeled using Tait's equation of state for water, surrounded by a ''membrane'' with surface tension that increased linearly with areal expansion. The effect of a nearby ''cell'' on bubble response, the resulting pressure field around the ''cell,'' and ''cell'' membrane tensions were investigated. Preliminary results were compared with critical values for hemolysis reported in the literature. [Work supported by NIH.]

3:00–3:20 Break

3:20

4pBB9. Model for interaction of bubbles in a cloud near a rigid surface. Evgenia A. Zabolotskaya, Yurii A. Ilinskii, G. Douglas Meegan (Appl. Res. Labs., Univ. of Texas, Austin, TX 78713-8029), and Mark F. Hamilton (Univ. of Texas, Austin, TX 78712-1063)

Bubble clouds produced during lithotripsy undergo complicated motions including bubble interactions that may inhibit kidney stone comminution. Our study of bubble interactions is motivated by high-speed photographs reported by Pishchalnikov *et al.* [J. Acoust. Soc. Am. **114**, 2386 (2003)]. In the work reported here, we simulated the observed bubble motion with a model based on the equations derived by Zabolotskaya [Sov. Phys. Acoust. **30**, 365 (1984)]. The equations for interaction of two bubbles were generalized and solved numerically for a cluster of n bubbles near a rigid boundary, which represents the stone. The initial spatial distribution of bubbles in three dimensions was assumed to be random. When a short negative pressure pulse was applied, the simulated bubbles grew in size. When two bubbles touched each other, they were merged into a single bubble that conserved mass of the gas. Results are presented in selected planes intersecting the bubble cloud for different instants of time. Bubble interaction was found to reduce the maximum sizes to which the bubbles grow. The bubbles near the rigid boundary are constrained by neighboring bubbles and grow less rapidly, and to smaller sizes, than other bubbles. Interactions within the cloud thus suppress bubble growth and cavitation. [Work supported by ARL:UT IR&D.]

3:35

4pBB10. Importance of pulse synchrony to stone comminution in dual-pulse lithotripsy. Yuri A. Pishchalnikov, James A. McAteer (Dept. of Anatomy and Cell Biol., Indiana Univ., 635 Barnhill Dr., Indianapolis, IN 46202, yura@anatomy.iupui.edu), Irina V. Pishchalnikova, Richard J. VonDerHaar, James C. Williams, Jr., Andrew P. Evan (Indiana Univ., Indianapolis, IN 46202), Michael R. Bailey (Ctr. for Industrial and Medical Ultrasound, Seattle, WA 98105), and Robin O. Cleveland (Boston Univ., Boston, MA 02215)

We have characterized the acoustic output and *in vitro* stone breakage of the first dual-pulse lithotripter approved for patient treatment. Direx Medical Systems provided their Duet lithotripter and two timing circuits—a standard circuit that fires both electrodes in synchrony and a modified circuit that allows pulse delay. We assessed the effect of synchronous and ~10-μs delayed dual pulses on breakage of model stones. When synchrony was very close the waveform had a single peak ~70 MPa P^+, followed by a negative trough ~12 MPa The waveform from one head was about half this amplitude. Pulses close in timing but not synchronous (~1–2-μs apart) showed amplitude-summed conjoined peaks. Dual pulses fired ~10 μs apart appeared as independent SWs. Simultaneous pulses broke stones better than pulses delivered by only one head (p <0.001) and better than ~10-μs delayed pulses (p<0.0001). Delayed dual pulses were also less efficient than pulses delivered by one shock head (p<0.001). Pulse timing is critical to stone breakage in dual-pulse lithotripsy. When pulse timing is very close, breakage is better than pulses from one source. The observation that asynchronous pulses are less efficient highlights the importance of precision in dual-pulse SW sources. [Work supported by NIH-DK43881, ONRIFO-N00014-04-1-4010.]

3:50

4pBB11. The role of shear and longitudinal waves in the kidney stone comminution by a lithotripter shock pulse. Oleg A. Sapozhnikov (Dept. of Acoust., Faculty of Phys., M.V. Lomonosov Moscow State Univ., Leninskie Gory, Moscow 119992, Russia, oleg@acs366.phys.msu.ru), Robin O. Cleveland (Boston Univ., Boston, MA 02215), and Michael R. Bailey (Univ. of Washington, Seattle, WA 98105)

Shock wave lithotripsy has been in clinical use for 20 years but there is no consensus as to the main mechanism of kidney stone comminution. Experiments show that several mechanisms might be involved, including cavitation, spallation, and dynamic fatigue. Until recently, little attention was paid to shear elasticity of the stone material, i.e., mechanical load was mainly attributed to the longitudinal waves. In a previous numerical study, we found that shear elasticity resulted in tremendous change in the stress pattern inside cylindrical stones. The numerical model has been extended to study elastic waves in asymmetric inhomogeneous stones. Strains and stresses in the stone are calculated based on the Lamé equation for an isotropic elastic medium. Lithotripter shock waves of various temporal and spatial profiles were considered according to several clinical models of lithotripters. Maximum compression, tensile and shear stresses are predicted as a function of stone dimension and shape. The model predicts that both shear and longitudinal waves play an important role in creating the regions of excess stresses where cracks can be formed. The results of modeling are compared with the experimental observations. [Work supported by ONRIFO, CRDF, NIH-Fogarty, RFBR, NIH, and Whitaker Foundation.]

4:05

4pBB12. Comparison between an open-cage electrode lithotripter and an encapsulated electrode lithotripter. Parag V. Chitnis and Robin O. Cleveland (Aerosp. and Mech. Eng. Dept., Boston Univ., 110 Cummington St., Boston, MA 02215, pchitnis@bu.edu)

We compare the acoustic and cavitation field of two electrohydraulic lithotripters: one employed an open-cage electrode (OCE) and the other an encapsulated electrode (ECE). Acoustic pressure was measured with a fiberoptic probe hydrophone and cavitation using a passive detector consisting of two confocal transducers. The focused shock waves (FSW) of the two lithotripters were similar in shape and at 20 kV, peak positive pressure ($P+$) for the OCE was 23.2±4.4 MPa and peak negative pressure ($P-$) was 9.0±1.5 MPa. For the ECE the peak pressures were 30.8±9.4 and −8.2±1.3 MPa, respectively. However, the direct waves (DW) were different. The OCE-DW was triangular shaped and almost entirely positive pressure, whereas the ECE-DW consisted of numerous cycles and had a peak negative pressure of 1.3 MPa. For both electrodes the passive cavitation measurements exhibited two acoustic emissions characteristic of inertial collapses. For the OCE the first emission was weak (in comparison to the second emission) and was attributed to scattering of the FSW. For the ECE both emissions were of similar amplitude, indicating that the DW generated cavitation bubbles that were forced to a violent collapse by the FSW. The ECE produces two violent cavitation events for every shock wave. [Work supported by HMT and NIH.]

4:20

4pBB13. Volume measurements of the cell destruction zone and thermal decomposition zone in the focus of a shock wave transducer. Igor Mastikhin (MRI Ctr., Phys. Dept, Univ. of NB, 8 Bailey Dr., Fredericton, NB E3B 5A3, Canada, mast@unb.ca), Vyacheslav Teslenko (Lavrentiev Inst. of Hydrodynamics, Novosibirsk 630090, Russia), and Valery Nikolin (Inst. of Cytology and Genetics, Novosibirsk, 630090 Russia)

Evaluation of the volume of the cell destruction zone is of interest in biomedical applications of shock waves (SW). The volume depends on mechanical properties of the cell membranes and is different for different cell types. In this work, we evaluated the cell destruction volume for two different cell types, tumor cells Crebs-2 and red blood cells. We used 0 70 0.5-s SW pulses with 45-MPa pressure in the focal zone. The concentration of destroyed cells was counted by dyeing in the case of tumor cells, and by spectrometry of released hemoglobin in the case of RBC. The cell destruction volume was calculated from destruction versus pulse number data and measured as 0.0135 ml for tumor cells. For RBC, the volume was 0.021 ml. To evaluate the effective volume of thermal zone, we used EPR signal of stable disulphide biradicals. Under SW action, S–S bonds of the biradicals rupture. The volume measurements were 0.003 ml. Since for that biradical, S–S bonds rupture at temperatures >80°C, and concentration of free radicals was an order lower (measured by spin traps) than of the produced monoradical, the rupture was caused by thermal decomposition. Thermal effects can play a significant role in SW action.

Session 4pED

Education in Acoustics: Careers in Acoustics

Uwe J. Hansen, Chair
Physics Department, Indiana State University, Terre Haute, Indiana 47809

Chair's Introduction—2:00

Invited Papers

2:05

4pED1. Architectural-acoustics consulting. Anthony K. Hoover (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776)

Consulting involves both the science of acoustics and the art of communication, requiring an array of inherent and created skills. Perhaps because consulting on architectural acoustics is a relatively new field, there is a remarkable variety of career paths, all influenced by education, interest, and experience. Many consultants juggle dozens of chargeable projects at a time, not to mention proposals, seminars, teaching, articles, business concerns, and professional-society activities. This paper will discuss various aspects of career paths, projects, and clients as they relate to architectural-acoustics consulting. The intended emphasis will be considerations for those who may be interested in such a career, noting that consultants generally seem to thrive on the numerous challenges.

2:20

4pED2. Acoustical consulting—Reflections on a challenging career. David Braslau (David Braslau Assoc., Inc., 1313 5th St. SE, Ste. 322, Minneapolis, MN 55414)

The acoustical consulting profession can be entered in a number of ways. The most direct approach is to obtain a degree in acoustics and join a large consulting firm immediately after graduation. Acoustical consulting can also be entered indirectly from various fields of engineering or physics which can provide a somewhat broader background. These disciplines might include, for example, structural engineering and structural dynamics, mechanics of materials, dynamic behavior of solids or geophysics. Acoustical consulting specialization can be very broad or very narrow as seen from the National Council of Acoustical Consultants capability listing. As an acoustical consultant, one must address a wide range of problems which provides both the challenges and joys of this profession. Technical capabilities and professional judgment are constantly developed from exposure to these problems and through interaction with other members of the profession. Selected case studies including sound isolation in buildings, noise and vibration from blasting, control of noise from environmental sources, acoustical design of classrooms and performing spaces, and product design demonstrate the variety of challenges faced by an acoustical consultant.

2:35

4pED3. What happens when you want to talk to the animals: lessons of a career in animal bioacoustics. Ann E. Bowles (Hubbs-SeaWorld Res. Inst., 2595 Ingraham St., San Diego, CA 92109, annb1@san.rr.com)

Animal bioacoustics (AB) is the study of sound in nonhuman animal biology. I entered the field because I was interested in the evolution of language, and I wanted to study the acoustic communication of whales and dolphins. Topics like this within the scope of AB make the discipline accessible to students and laypeople. Although career opportunities are limited (professionals declaring AB as their primary area represent only 3% of ASA membership [http://www.acoustics.org/WIA, statistics for 2000]), an interest in AB can foster entry into more marketable disciplines. It has been a particularly important avenue for bringing women into careers in traditionally male-dominated subject areas. For example, women represent 14% of the ASA membership and 12% or fewer of those declaring Underwater Acoustics, Engineering Acoustics, and Noise as their primary interest. However, 25% of those declaring AB as their primary interest are women, and AB includes all three topic areas within its scope. Unfortunately, AB is still fairly inaccessible to interested lay professionals such as educators, science writers, and environmental planners. By helping them to develop a deeper understanding of topics in AB, ASA can help them foster careers in acoustics.

2:50

4pED4. Biomedical engineering: A platform for research and innovation in ultrasound. Christy K. Holland (Dept. of Biomed. Eng., MSB 6167, 231 Albert Sabin Way, Cincinnati, OH 45267-0586)

An undergraduate or graduate degree in biomedical engineering prepares students to solve problems at the interface between engineering and medicine. Biomedical engineering encompasses evolving areas such as advanced medical imaging for diagnosis and treatment of disease, tissue engineering for designing and manufacturing biological implants for damaged or diseased tissues and organs, and bioinformatics for determining which genes play a major role in health and disease. Biomedical engineering academic programs produce graduates with the ability to pursue successful careers in the biomedical device industry or to obtain advanced degrees leading to careers in biomedical engineering research, medicine, law or business. Biomedical engineering majors take courses in biology, anatomy, physics, chemistry, engineering, mathematics and medical product design and value life-long learning. Students

learn to work effectively in interdisciplinary teams comprised of individuals with diverse social, cultural and technical backgrounds. Biomedical engineering is becoming increasingly important in imaging and image-guided research. Some examples of innovative ultrasound technology under development are ultrasound devices to accelerate the dissolution of blood clots, advanced surgical instruments with ultrasound guidance and ultrasound contrast agents for targeted drug delivery. Biomedical engineering is a great career choice for technically minded individuals who endeavor to work on applied problems that are medically relevant.

3:05

4pED5. Life as an acoustician in industry, academia, and government service. Mardi C. Hastings (ONR, 800 N. Quincy St., Arlington, VA 22217, mardi_hastings@onr.navy.mil)

Acoustics is a science that has very broad applications, which affect all different areas of our lives. During the last 20 years, I have combined family with a career as an acoustics engineer in industry, a tenured faculty member at a university and, most recently, a program manager in a government agency. In these positions I have worked in several areas of acoustics, including noise control, structural acoustics, building acoustics, sound quality, physical acoustics, acoustic materials, underwater acoustics, biomedical ultrasound, physiological acoustics, and bioacoustics. Although the fundamental science of sound is the foundation of all these areas, communication of ideas, problems, and solutions varies greatly from industry to academia to government. Thus knowing the science and how to use it are not enough, as communication skills and the ability to adapt them to changing environments are essential for a successful career. In addition to describing life as an acoustician in industry, academia, and government service, I will present several examples of how even though the acoustic fundamentals are the same, how they are communicated could become a disaster or save the day.

3:20

4pED6. Acoustics careers for engineers. Uwe J. Hansen (Indiana State Univ., Terre Haute, IN 47809)

Many acoustics opportunities in industry, government laboratories, and academics rely on a background in mechanical engineering, electrical engineering, or physics. Acoustics deals principally with generation, propagation, and perception of sound. Engineering application, include among many other things, the study and control of structural vibrations, machinery analysis and maintenance, and industrial noise control. Thus, for example, the aircraft industry employs engineers to study vibrational characteristics of items such as turbine blades or entire fuselage assemblies. Among the many techniques utilized are holographic interferometry and modal analysis. Some of these methods will be illustrated.

3:35–3:45 Break

THURSDAY AFTERNOON, 27 MAY 2004 CONFERENCE ROOM K, 1:15 TO 4:55 P.M.

Session 4pMU

Musical Acoustics: New Research on Pre-1929 Instruments

Thomas D. Rossing, Cochair
Physics Department, Northern Illinois University, De Kalb, Illinois 60115

D. Murray Campbell, Cochair
Department of Physics and Astronomy, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom

Chair's Introduction—1:15

Invited Papers

1:20

4pMU1. Pitch bending on the cornet. D. Murray Campbell (School of Phys., Univ. of Edinburgh, Mayfield Rd., Edinburgh EH9 3JZ, UK, d.m.campbell@ed.ac.uk)

A successful model of a brass wind instrument must be able to describe the playing technique known as "lipping," in which the player can bend the pitch of a note by modifying the setting and tension of the lips. This technique is of special importance on the lip-reed instrument known as the cornet or cornetto, which was important in the 15th and 16th centuries but fell out of use by the end of the 18th century. The instrument consists of a short wooden tube of approximately conical bore with seven side holes. Chromatic notes must be obtained by cross fingering, and good intonation requires fine control of the lipping technique. Recent experimental studies of lipping on the cornet are described, and implications for modeling of the lip reed are discussed.

1:40

4pMU2. The acoustical engineering of brasswind instruments 1779–1929. Arnold Myers (Univ. of Edinburgh, Reid Concert Hall, Bristo Square, Edinburgh EH8 9AG, UK, A.Myers@ed.ac.uk)

At the start of this 150-year period, brass musical instruments were made to traditional designs, which were developed by trial and error, the fittest surviving. Mechanical inventions (most importantly the valve) greatly widened the possibilities for bore engineering. Increasingly through this period, instrument designers were influenced by the developing science of acoustics. By 1929 most of the range of instruments in use today had been developed and acoustical tools were in use in optimizing the design of instruments. Considering the factors of greatest importance in determining the acoustical response of a brass instrument to be bore profile, bell flare (cutoff frequency), and mouthpiece geometry, landmarks in the development of existing instrumental types and the creation of new models are surveyed. The contributions of Stoelzel, C.M. Pace, Sax, Bayley, Blaikley, Webster, and Couturier are discussed. This paper is based on research involving direct examination of several hundred instruments from 1779–1929 located in museums worldwide.

2:00

4pMU3. Acoustics of old Asian bells. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

The art of casting bronze bells developed to a high level of sophistication in China during the Shang dynasty (1766–1123 BC). Many chimes of two-tone bells remain from the Western and Eastern Zhou dynasties (1122–249 BC). With the spread of Buddhism from the third century, large round temple bells developed in China and later in Korea, Japan, and other Asian countries. Vibrational modes of some of these bells have been studied by means of holographic interferometry and experimental modal testing. Their musical as well as acoustical properties are discussed.

2:20

4pMU4. The design and analysis of new musical bells. Neil M. McLachlan (School of Aerosp., Manufacturing and Mech. Eng., RMIT, GPO Box 2476V, Melbourne Vic. 3000, Australia, neil.mclachlan@rmit.edu.au)

The design and analysis of a series of new musical bells will be presented in this paper. Modal analysis of a wide range of bell-like geometries using FEA revealed the presence and significance of transverse axial modes in unconstrained bell models, leading to a new understanding of the relationships between bell geometry and modal behavior. This understanding was used to adjust simple parametric models of bell geometry to arrive at appropriate geometries to begin numerical shape optimization for the design of bells with a range of desired overtone tunings. Pitch salience is well known to depend on the degree of harmonic relationships between pure tones in complex stimuli. Bells intended to produce a single, highly salient pitch were designed and manufactured with up to the first 7 overtones tuned to the harmonic series. Other bells with overtones tuned to subsets of two or three harmonic series were also designed and manufactured. These bells were intended to produce multiple pitch perceptions of approximately equal strength. Spectral analysis and range of numerical psycho-acoustic models are used to evaluate the sounds of manufactured bells against these design objectives. [I would like to acknowledge the close collaboration of Dr. Anton Hasell of Australian Bell.]

2:40

4pMU5. Normal modes of different types of mandolins. David J. Cohen (Cohen Musical Instruments, 9402 Belfort Rd., Richmond, VA 23229) and Thomas D. Rossing (Northern Illinois Univ., DeKalb, IL 60115)

The vibrational modes and sound spectra of some pre-1929 archtop mandolins and pre-1922 Neapolitan (“bowlback”) mandolins have been studied. The results have been compared with those obtained previously on archtop mandolins constructed more recently [D. Cohen and T. D. Rossing, CASJ **4**(2), 48–54 (2000), D. Cohen and T. D. Rossing, Acoust. Sci. Tech. **24**, 1–6 (2003)]. Some obvious and predictable differences between the Neapolitan mandolins and the archtop mandolins were found. The very stiff bowls of the Neapolitans do not contribute to corpus vibrations below about 1.2 kHz. The ladder-braced top plates of the Neapolitans are also quite stiff, with the (0,0) mode first occurring at or above 500 Hz. The (0,0) modes in archtop mandolins generally occur at lower frequencies. Archtop mandolins with f-holes generally have either longitudinal bracing or X-bracing, with the results that the modes involving cross-grain bending [e.g., (1,0), (2,0), etc.] occur at lower frequencies than the modes involving bending along the grain [e.g., (0,1), (0,2), etc.]. In the ladder-braced Neapolitans, the modes involving cross-grain bending occur at higher frequencies than the modes involving bending along the grain.

3:00–3:10 Break

3:10

4pMU6. The acoustics of carved Baltic psaltery. Andres Peekna (Innovative Mech., Inc., 265 Coe Rd., Clarendon Hills, IL 60514-10299, innmech@comcast.net) and Thomas Rossing (Northern Illinois Univ., DeKalb, IL 60115)

The Baltic psaltery family of plucked string instruments includes the kantele (Finland), the kannel (Estonia), the kokle (Latvia), the kankles (Lithuania), and the wing-shaped gusli (Northwestern Russia). In its archaic, carved form, it has a limited range, 5–13 strings, usually tuned diatonically. By means of electronic TV holography, we studied the modes of vibration of several psalteries based on historic instruments. On the better instruments, the main body resonances are well distributed in frequency so that they support the various strings. Good string-to-soundbox coupling also appears to play a role. A useful method for studying string-to-soundbox coupling involves scanning at intervals as low as 0.1 Hz for narrow peaks within the nominal tuning range of the strings,

and comparing them to their neighboring body resonances, while using electronic TV holography. Predictions of the Helmholtz resonance from sound-hole dimensions and air-cavity volume while neglecting damping in the sound holes yield upper limits when many small sound holes are involved. The locations of the sound holes, as well as their area, are found to have significant effects on sound quality and volume.

3:30

4pMU7. The harpsichord in 1929 and the emperor's new clothes. Edward Kottick (Univ. of Iowa, Iowa City, IA 52242)

By the end of the 19th century the harpsichord was considered obsolete. Viewed from the vantage point of a century of brilliant advances in piano technology, it was an outworn, outmoded relic of a distasteful past. It was overly delicate in nature; its soundboard was too thin, and thus unstable; its case construction was unsubstantial. Its light, unbushed keyboards were inadequate. Its quaint bird-quill plectra were viewed with disdain, and its sound was thought to be weak and unattractive. Nevertheless, by the year 1929 the harpsichord was enjoying a revival that was at that time in its 40th year, and that revival continues today. If this instrument was considered so inadequate, why has its revival been so successful and long-lasting? All the answers boil down to one word: sound. This paper will explore the contrast of the tonal qualities of the revival harpsichords and the antiques, and the way those qualities changed in response to the shifting tastes of builders, performers, and listeners. Ultimately, changing tastes led to a return to the sound of the classical harpsichord. Embedded in this tale is an important lesson on the dangers of characterizing the tone of any instrument.

3:50

4pMU8. Physical modeling of Tibetan bowls. Jose Antunes (Instituto Tecnologico e Nuclear, Appl. Dynam. Lab., 2686 Sacavem codex, Portugal, jantunes@itn.mces.pt) and Octavio Inacio (Instituto Politcnico do Porto, 4000-045 Porto, Portugal)

Tibetan bowls produce rich penetrating sounds, used in musical contexts and to induce a state of relaxation for meditation or therapy purposes. To understand the dynamics of these instruments under impact and rubbing excitation, we developed a simulation method based on the modal approach, following our previous papers on physical modeling of plucked/bowed strings and impacted/bowed bars. This technique is based on a compact representation of the system dynamics, in terms of the unconstrained bowl modes. Nonlinear contact/friction interaction forces, between the exciter (puja) and the bowl, are computed at each time step and projected on the bowl modal basis, followed by step integration of the modal equations. We explore the behavior of two different-sized bowls, for extensive ranges of excitation conditions (contact/friction parameters, normal force, and tangential puja velocity). Numerical results and experiments show that various self-excited motions may arise depending on the playing conditions and, mainly, on the contact/friction interaction parameters. Indeed, triggering of a given bowl modal frequency mainly depends on the puja material. Computed animations and experiments demonstrate that self-excited modes spin, following the puja motion. Accordingly, the sensed pressure field pulsates, with frequency controlled by the puja spinning velocity and the spatial pattern of the singing mode.

Contributed Papers

4:10

4pMU9. The acoustics of the Bagana. Stephanie Weisser (Universite libre de Bruxelles, CP 175, 50 av. F. Roosevelt, 1050 Brussels, Belgium) and Didier Demolin (Universidade de Sao Paulo, 01060-970 Sao Paulo, Brasil)

The Bagana lyre of Ethiopia is an instrument with a very particular timbre whose acoustic characteristics is determined by U-like leather strips placed between each string and the bridge. These leather strips give a deep buzzing sound that influences pitch duration, intenisty, and timbre. The sounds of the instruments were recorded with and without buzzers for comparison. Data were analyzed to characterize the main acoustic features of the instrument. In addition a high-speed camera (400 images/second) was used to describe the different vibratory modes. The main observation is that buzzers enhance the spectrum up to more than 10 kHz; they also modify the attack and the release of the sound. The duration and the loudness of the sound are increased because they enhance the energy's repartition in the spectrum, especially in the area around 1500–300 Hz. Another effect of the buzzers is that the instrument produces low pitch and loud sounds without having a big resonator or tall harmonic strings.

4:25

4pMU10. On the development of German beating-reed organ pipes during the 19th century. Jonas Braasch (Faculty of Music, McGill Univ., 555 Sherbrooke St. W., Montreal, QC H3A 1B9, Canada, braasch@music.mcgill.ca)

In the 19th century organ literature, it is often claimed that German organ builders generally adapted the way of building their beating-reed pipes after being influenced by new developments from England and France. To investigate whether this hypothesis is true or false, the reed-pipe sounds of several German historic organs and an English organ by Henry Willis were measured and analyzed. The outcome of the analysis, however, cannot confirm the given hypothesis. Organ builders of the 18th century, such as Gottfried Silbermann for example, were already able to build beating-reed pipes similar in sound to the pipes that are used nowadays in Germany. It is noteworthy that Silbermann used closed shallots in some of his stops, although they are thought to be one of the main inventions in the English and French organ reforms. The use of higher wind pressures, which is also a main part of this reform, on the other hand, never became a common standard in Germany, as was the case for France and Great Britain.

4:40

4pMU11. Acoustics of glass harmonicas. Thomas D. Rossing (Phys. Dept., Northern Illinois Univ., DeKalb, IL 60115)

Glass musical instruments are probably as old as glassmaking. At least as early as the 17th century it was discovered that wine glasses, when rubbed with a wet finger, produced a musical tone. A collection of glasses played in this manner is called a glass harp. Another type of glass harmonica, called the armonica by its inventor Benjamin Franklin, employs glass bowls or cups turned by a horizontal axle, so the performer need only touch the rim of the bowls as they rotate to set them into vibration. We discuss the modes of vibration of both types of glass harmonica, and describe the different sounds that are emitted by rubbing, tapping, or bowing them. Rubbing with a wet finger tends to excite only the (2,0) mode and its harmonics through a ''stick-slip'' process, while tapping excites the other modes as well.

Session 4pNSa

Noise and Architectural Acoustics: Distinguished Lecture "Noise: My 62 Years of It''

Leo L. Beranek, Chair
975 Memorial Drive, Suite 804, Cambridge, Massachusetts 02138-5755

Chair's Introduction—1:15

Invited Paper

1:20

4pNSa1. Noise: My 62 years of it! Laymon N. Miller (1504 Harbor Court, Fort Myers, FL 33908-1651, laymnluce@aol.com)

Imagine getting paid for having fun! Well, in retrospect, it was fun; but there were several tough challenges. Even those are worth remembering. From 1941 to 1982, there were acoustic torpedoes, HVAC acoustics, noise and vibration in auditoriums, aircraft and airport noise, OSHA and industrial noise control, power plants, community noise problems, vbration, railroad and subway vibration control, legal acoustics, noise manuals, and noise courses—and a few other things that don't fit into those neat categories. Some specific jobs could be named, but that would take away the suspense and the surprise. But 1941 to 1982 is only 41 years. How about the other 20-odd years?

Session 4pNSb

Noise and Psychological and Physiological Acoustics: Noise and Society

Nancy Nadler, Chair
League for the Hard of Hearing, 50 Broadway, New York, New York 10004

Chair's Introduction—2:30

Invited Papers

2:35

4pNSb1. The nature of noise in society. Leslie Blomberg (Noise Pollution Clearinghouse, Box 1137, Montpelier, VT 05601)

Noise is unique among pollutants. It is, for example, the only pollutant commonly defined in subjective, psychological terms: ''unwanted sound.'' Ironically, noise experts work almost entirely in objective measures-sound pressure, sound power, Leq, Ldn, etc. This paper suggests that between the science and engineering of acoustics and the psychology of sound perception, in the societal and social context, lies the true nature of noise–its causes and effects. Enriched by contacts with tens of thousands of individuals over the past 8 years, the Noise Pollution Clearinghouse is working to redefine noise in societal terms. In understanding the nature of noise, the concepts of civility, sovereignty, environmental quality, quality of life, and connectedness to others are much more helpful than ''unwanted sound.''

2:55

4pNSb2. Noise pollution: A threat to our mental and physical well-being. Arline L. Bronzaft (Dept. of Psych., Lehman College, Bronx, NY 10468, albtor@aol.com)

While noise may not yet be in the forefront of the environmental movement, it is being recognized worldwide as a major environmental pollutant. In New York City, noise is the number one quality of life complaint, far outweighing other quality of life complaints, and throughout the United States it has been noted as a major reason for people moving from their homes. Although there is a need for additional research to confirm the health/noise link, the World Health Organization has already recognized noise pollution as a serious health issue. There is certainly sufficient research to warrant warnings that noise is injurious to mental and physical health. Yet, despite this growing body of literature attesting to the relationship between noise and health impacts, government bodies have not yet invested the dollars needed to abate noise nor to educate people to the dangers of noise. Organizations such as the League for the Hard of Hearing, the Noise Pollution Clearinghouse, and the United Kingdom Noise Network have assumed the tasks of educating the public to the harmful effects of noise, of advocating anti-noise measures, and of urging public officials to move more assertively in lessening the din of our ever-increasing noisy society.

3:15

4pNSb3. Children's quality of life in a noisy world. Brigitte Schulte-Fortkamp (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany, brigitte.schulte-fortkamp@tu-berlin.de)

On 30 April 2003 during the International Forum Noise Awareness Day European experts from the fields of medicine, acoustics, sociology, psychology, city planning, and traffic regulation led the current discussion on the risks of noise pollution for children at the Institute of Technical Acoustics at the Berlin Technical University, Germany. Studies probing into the negative effects of noise exposure upon the psychic, cognitive, and emotional functions of children are of a quite recent date. Even the quality of life of children as auto-directed experience and functionality is only now starting to get recognized; of special interest are the changes in self-awareness, performance, and health. The Forum focused upon methodological reflections for the collection of data as well as studies regarding the living conditions of children under sound exposure from different perspectives. In this context the research of sound effects has taken initial tentative steps to break away from the tried-and-true procedures of the last 30 years. Next to applied science, investigating fundamentals will need to be considered. Aiming at a networking process is one of the goals conducting the ''Tag gegen Lrm—International Noise Awareness Day.''

3:35

4pNSb4. Training as a critical component of successful noise enforcement programs. Eric Zwerling (Rutgers Univ. Noise Tech. Assistance Ctr., 14 College Farm Rd., New Brunswick, NJ 08901)

The point of application of any noise enforcement program is the enforcement officer. The quality of their training is of paramount importance in determining their efficacy in resolving complaints in the field or, failing that, in court. Some of the critical components that must be addressed in a training program are the technology, techniques and strategies of legally valid sound level measurement; documentation of measurement parameters and results; calculation of corrected source sound levels; managing the expectations of complainants; negotiations with alleged violators; and compliance determination methods for nonmetered performance standards. A strong emphasis must be on practical field measurements. The training must assist the enforcement officer to become comfortable with the process, motivating the officer to embrace the new skill, rather than resenting a new task. It is important to take into account the background of the students, professionally, and as individuals, as well as the institutional culture of their agency. The better prepared an officer is to go to court, the less likely is that possibility. A well designed and executed program, represented by its field officers, provides significant deterrence. Thirteen years of training experience at the Rutgers Noise Technical Assistance Center is reviewed.

3:55

4pNSb5. The effects of removing barriers to hearing protector use. Mark Stephenson (Ctr. for Disease Control and Prevention/ NIOSH, Cincinnati, OH)

When high level noise exposure is unavoidable, wearing hearing protectors is essential to preventing hearing loss and preserving the quality of ones life. Unfortunately, many people fail to wear hearing protectors when they are exposed to loud noise. Focus groups and surveys conducted by the National Institute for Occupational Safety and Health (NIOSH) have identified a set of barriers frequently cited as reasons for not wearing hearing protection. NIOSH has employed these data to develop a field study designed to address barriers to hearing protector use. Interim results have demonstrated that by applying contemporary health communication theory to focus training messages and methods on specific barriers, it is possible to positively influence attitudes and behaviors associated with hearing protector use. This paper will describe the methods used in this study and discuss the results obtained to date.

4:15

4pNSb6. Hearing loss and tinnitus in adolescents and young adults. Alice Holmes (Univ. of Florida, Box 100174, Gainesville, FL 32610, aholmes@hp.ufl.edu)

Little attention has been paid to hearing abilities and the effects of noise on the normal adolescent and young adult population. A series of studies will be presented on the prevalence of hearing loss and reported effects of hearing loss and tinnitus in adolescents and young adults from different cultural backgrounds. Adolescents and young adults from different backgrounds may tend to seek or avoid various noise environments that could be detrimental to their hearing and cause tinnitus. Attitudes and exposures to noise environments were evaluated to see if these may be correlated with their hearing losses and/or tinnitus. In addition, these adolescent and young adult subjects reported how often they used hearing protection in various noise environments. Finally, the issues of quality of life and the need for hearing conservation programs with these populations will be presented.

4:35

4pNSb7. Regulations against the human nature. Fernando J. Elizondo-Garza (Acoust. Lab., FIME, Univ. A. de Nuevo Leon, P.O. Box 28 F, Cd. Universitaria, San Nicolas, 66450, N.L., Mexico, fjelizon@ccr.dsi.uanl.mx)

The discussion around the concept of the addiction to noise has evidenced the importance of noise for the human being and explains why in some cases the regulations fail to control the noise in cities. In this presentation the different uses, consciously or unconsciously, of the noise will be analyzed, uses that go from habits to maybe addictions. Also discussed are the implications of establishing regulations against the human nature as well as the importance of education to manage the noise and design acoustically instead of trying to ban the noise in some social circumstances.

4:50

4pNSb8. Data treatment method to improve the determination of indexes related to the directive 2002/49/EC about environmental noise management. Jose Luis Cueto, Silvia Rivas, and Ricardo Hernandez (Laboratorio de Acustica y Vibraciones, Univ. of Cadiz, C.A.S.E.M. Polgono Ro San Pedro s/n 11500 Puerto Real. Cdiz., Spain, joseluis.cueto@uca.es)

The European Directive 2002/49/EC defines useful procedures in order to evaluate, by means of indexes, the acoustic pollution affecting urban areas. To apply those methods correctly, it is necessary to get a properly designed acoustic database. Usually the environmental information managed by the laboratories shows a different kind of problem. Nowadays the Acoustic Laboratory of the University of Cadiz (Spain) is very interested in the developing of new data-processing tools. These tools overcome the errors in the calculation of the indexes due to the use of inappropriate databases. We focus not only on the determination of the possible problems, but on the best solutions. We suggest in this paper a simple data treatment method that permits one to reach more accurate indexes in spite of such common problems.

THURSDAY AFTERNOON, 27 MAY 2004 CONFERENCE ROOM L, 1:00 TO 4:30 P.M.

Session 4pPA

Physical Acoustics: Scattering

Ronald Kumon, Chair
Department of Physics, University of Windsor, 401 Sunset Avenue, Windsor, Ontario N9B 3P4, Canada

Contributed Papers

1:00

4pPA1. The Schoch effect as a means to detect the principal axes of a piezoelectric crystal. Nico F. Declercq, A. Teklu, M. A. Breazeale (Natl. Ctr. for Physical Acoust., The Univ. of Mississippi, Oxford, MS 38677, NicoF.Declercq@Ugent.be), Joris Degrieck (Ghent Univ., B-9000 Ghent, Belgium), and Oswald Leroy (KU Leuven Campus Kortrijk, B-8500 Kortrijk, Belgium)

When a bounded beam is incident on a piezoelectric crystal at the right angle, surface waves are generated which result in the Schoch effect. This effect consists of the appearance of two reflected beams instead of one, with a null strip in between. The characteristics of this Schoch effect are determined by the physical properties of the solid under consideration, but also by the direction of the principal axes. Therefore, the effect can be used to determine those directions. A comparison of theory and experiments on lithium niobate is shown. This method can be used as a practical tool for determining the principle axes, since these are not visible to the eye. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.) and by a NATO Collaborative Linkage Grant.]

1:15

4pPA2. Air acoustic propagation in a multiple scattering environment: Application to target detection in a forest. W. J. Higley, Philippe Roux, Jit Sarkar, and W. A. Kuperman (Marine Physical Lab., Scripps Inst. of Oceanogr., NTV Bldg. 4th Fl., 8820 Shellback Way, La Jolla, CA 92093-0238, whigley@ucsd.edu)

A small-scale acoustic experiment in a multiple scattering medium has been conducted in the audible bandwidth of 4 to 6 kHz. Approximately 200 1-m-long PVC pipes were stood upright in a 8×10 m square between an array of eight microphones and a parallel array of ten speakers to simulate conditions in a dense forest. The Green's functions of the propagation between each microphone and speaker were measured several times with no changes to the medium. Then, an additional, moving scatterer was introduced into the medium and the Green's functions were again acquired. An algorithm has been developed to recognize the difference in the propagation when the additional scatterer is present and when it is not. Further analysis of this data shows the effective cross-section of the additional scatterer. A potential use of such an algorithm and system is the detection and classification of foreign scatterers in a forest.

1:30

4pPA3. Multiple Bragg scattering in arbitrary periodic acoustic waveguides. Dan Valente (Grad. Prog. in Acoust., Penn State Univ., P.O. Box 30, State College, PA 16804, dpv110@psu.edu) and David C. Swanson (Penn State Univ., State College, PA 16804)

Periodic acoustic waveguides allow the propagation of only certain bands of frequencies, thus making them very effective acoustic filters. The particular distribution of these bands depends critically on the waveguide geometry. If the variation of the waveguide walls happens to be sinusoidal, then the waveguide preferentially reflects only the frequency which matches the Bragg condition for complete backscatter. The reflection of a single frequency in the sinusoidal waveguide seems to contrast the case of a more general periodic structure, where numerous bands of frequencies are disallowed. In this work, the relationship between the sinusoidal waveguide and an arbitrary periodic waveguide is studied using a transmission line model. Inspection of the input reflection coefficient as a function of frequency yields striking similarities to the Fourier transform of the wall variation, implying that a waveguide whose walls are defined by an arbitrary periodic function independently Bragg reflects the individual Fourier

components of the wall variation function. The specific example of a waveguide with a square wave variation (a waveguide with periodic expansion chambers) will be discussed and compared to the literature, as well as the possibilities of recovering the waveguide geometry from the input reflection coefficient.

1:45

4pPA4. Scattering of elastic waves in textured media. Liyong Yang and Joseph A. Turner (Dept. of Eng. Mech., Univ. of Nebraska, W317.4 Nebraska Hall, Lincoln, NE 68588-0526, jaturner@unl.edu)

The investigation of wave propagation and scattering of elastic waves in heterogeneous, anisotropic media is of substantial interest to quantitative nondestructive evaluation and materials characterization. The scattering of elastic waves in polycrystalline media is primarily due to interaction with the grains. Knowledge of wave velocity and attenuation may be used to infer material texture in polycrystalline aggregates. In this presentation, a model for wave propagation and scattering in polycrystalline materials with texture is presented. Attenuations and wave velocities are discussed for a general orthorhombic material made up of cubic crystallites. The attenuations of each wave type are derived as a function of dimensionless frequency and wave propagation direction, respectively, for given orientation distribution coefficients (ODCs). The ODCs are, in essence, the coefficients of an expansion of crystallite of orientation distribution function (ODF) in terms of a series of generalized spherical harmonics. A relationship between the phase velocity and recrystallization variables, such as annealing time, is also investigated for specific examples. Finally, numerical results are presented and discussed in terms of the relevant dependent parameters. The results are anticipated to advance the field of materials characterization of statistically anisotropic media. [Work supported by DOE.]

2:00

4pPA5. Scattering of two-dimensional periodic gratings composed of cylindrical cavities in an elastic medium. Sebastien Robert, Hervé Franklin, and Jean-Marc Conoir (LAUE, Universit du Havre, place R. Schuman, 76610 Le Havre, France)

A theoretical and a numerical calculation of the scattering by a two-dimensional periodic grating composed of parallel cylindrical cavities embedded in an elastic medium are presented. The major point of the method is that the whole grating can be decomposed as a series of a finite or an infinite number of infinite rows of periodically spaced identical cavities. The scattering by each row is determined by an exact self-consistent multiple scattering calculation developed in the elastic case. The propagation of the waves from row to row can then be determined by an iterative method or from Bloch's theorem. Compared to previous works on gratings of elastic scatterers in a fluid, this study emphasizes new and interesting results. Among them are the systematic formation of new stopping and passing bands when the scattering is resonant, and the presence of frequency domains of higher coupling between the longitudinal and transverse waves, corresponding to the propagation of guided waves between two reticular planes of the grating. Examples of gratings with only one periodicity direction are also studied.

2:15

4pPA6. Backscattering in the vicinity of a mode cutoff. Alan M. Whitman (Villanova Univ., Villanova, PA 19085), Mark J. Beran, and Shimshon Frankenthal (Tel Aviv Univ., Ramat Aviv, Israel)

In a previous paper [Whitman *et al.*, Waves Random Media **13**, 269–286 (2003)] a set of coupled equations was derived that describes the intermodal scattering of acoustic radiation in a duct whose speed of sound varies randomly in space and time. In that paper the main interest was in modes that were not near cutoff. Here the solution of these equations in the vicinity of the cutoff is treated. It is found that near cutoff almost all of the energy is reflected back independently of the other duct parameters. In addition to presenting this result the mathematical structure of the equations in these regions is analyzed in order to elucidate the reason for the behavior. Some numerical results are also presented.

2:30

4pPA7. Backscattering in space- and time-dependent random media acoustic intensity fluctuations. Shimshon Frankenthal and Mark J. Beran (Faculty of Eng., Tel Aviv Univ., Ramat Aviv 69978, Israel, shim@eng.tau.ac.il)

In two recent publications [(1) S. Frankenthal and M. J. Beran, Waves Random Media **13**, 241–268 (2003); (2) A. M. Whitman, M. J. Beran, and S. Frankenthal, *ibid*. **13**, 269–286 (2003)], an incremental slab formulation was employed to derive the equations for the intensity of acoustic signals that undergo scattering in a medium whose sound speed fluctuates randomly both in space and time. These equations were used to track the propagation of a narrow-band pulse and its echoes across a one-dimensionally stratified slab, and to compute the effects of intermodal scattering on the modal intensities along a sinusoidally excited duct. Here, following a brief review of the main assumptions and the major features of the intensity equations, the same formulation is used to derive equations for the mean-square intensity, which is needed to calculate the intensity fluctuations. With sinusoidal excitation, an analytical solution of these equations is obtained for propagation across a one-dimensionally stratified slab, and a system of differential equations, which can be solved by the techniques used in Ref. 2, is derived for the modal mean square intensities in a duct.

2:45

4pPA8. Multilayer impedance coating for submerged objects—theory and applications. Ronald Hughes, Jan Niemiec, and Herbert Uberall (Naval Surface Warfare Ctr., Carderock Div., Bethesda, MD 20817-5700)

We consider acoustic returns from water-immersed objects with partially absorptive single or multiple coating layers. For sufficiently high absorptivity, the effect of the layers, even when applied to a strongly reflective substructure, can be described by an effective impedance Z, acting at the outer surface of the top layer. Expressions for Z may be obtained following procedures for the corresponding case of electromagnetic scattering by Uberall. In acoustics, the surface impedance of a one-layer coating was obtained by Brekhovskikh and is presented here in a very concise form. For two-layer coatings, Z is obtained analytically by a procedure used by Glegg for sediment layers. Equations by Folds and Loggins or Brekhovskikh can be used for a multilayer case to obtain numerical solutions. With these procedures for obtaining surface impedances, applications have been developed for the study of acoustic reflections from coated submerged structures.

3:00–3:15 Break

3:15

4pPA9. Multiple scattering from a random distribution of elastic cylinders in a fluid: Reflection and transmission coefficients of a slab. Pierre-Yves Le Bas, Francine Luppe, and Jean-Marc Conoir (LAUE Université du Havre, Pl. Robert Schuman, 76610 Le Havre, France, pierre-yves.lebas@univ-lehavre.fr)

The reflection and transmission coefficients of a slab containing a random distribution of elastic, infinitely long cylindrical scatterers in water are studied. The slab is immersed in water and insonified by a harmonic plane wave. All scatterers are identical, parallel, with axes normal to the incidence plane. A uniform random distribution is supposed. An extension of the Fikioris and Waterman model is developed in order to describe the propagation of the coherent wave in the slab. Its dispersion equation is numerically solved, providing its velocity and its attenuation. When there is only one solution of the dispersion equation that corresponds to an actually propagating wave, the slab is shown to behave as a dissipative equivalent fluid medium (effective medium). The results are compared to

those obtained from the Waterman and Truell model and to those of Foldy. The reflection and transmission coefficients are then written in a compact form similar to that of a fluid slab, so that the frequency dependence of the density of the effective medium may be studied. Results are shown for different concentrations of scatterers and for two different kinds of cylindrical scatterers.

3:30

4pPA10. Reflection coefficient of a water-loaded monoclinic plate: Obtaining of a factorized expression. Olivier Lenoir and Lionel Guénégou (LAUE UMR CNRS 6068, Univ. Le Havre, Pl. R. Schuman, 76610 Le Havre, France, lenoir@univ-lehavre.fr)

For an isotropic plate, it was shown by Schoch and used by Überall *et al.* that the reflection coefficient R of an elastic plate immersed in a fluid can be written in terms of the $C_{A,S}$ functions and the ratio τ of the acoustic impedances in the plate and in the fluid. The denominator of R is the product of the $(C_{A,S} \pm j\tau)$ functions. The roots of these functions correspond to the antisymmetric (A) and symmetric (S) vibration modes of the fluid-loaded plate. In this case, the Lamb-type guided waves are the combinations of dilatational and shear waves polarized in the incident plane. In this study, we consider an orthotropic plate where the incident plane is different from a symmetry plane; therefore, everything happens as if we dealt with a monoclinic plate. In this case, the guided waves are the sum of nonpure longitudinal waves, shear waves, and SH waves. Nevertheless, it can be shown that the expression of the reflection coefficient of the monoclinic plate has formally the same factorized form as the one of the isotropic plate. This expression is numerically compared to the one obtained by the stiffness matrix method developed by Rokhlin *et al.*

3:45

4pPA11. Self-organizing spatial pattern formation in wedge diffraction. Mitsuhiro Ueda (Dept. of IDE, Tokyo Inst. of Technol., O-okayama, Meguro-ku, Tokyo 152-8552, Japan, ueda@ide.titech.ac.jp)

It is well-known that the amplitude of potential near the apex of a 2D rigid wedge is proportional to the (PAI/WA)th power of the distance between the observation point and the apex, where PAI$=3.14$ and WA is an apex angle of the wedge measured in the free space. In the case of a semi-infinite plane (WA$=2$PAI), the square-root dependence is observed. This phenomenon should be regarded as a self-organizing spatial pattern formation, since this pattern arises for any waves incident to the apex. It has been, however, explained in terms of eigenfunctions, and no mechanism for this self-organization has been proposed so far. The boundary value problem for the potential on the boundary of the wedge is formulated using a new principle of diffraction, that is, the virtual discontinuity principle of diffraction (VDPD). By applying this formulation to observation points that are located near the apex, a relation that describes mutual dependence of the potential around the apex is obtained. It can be shown that this relation allows the stable solution only for the specific distance dependence. Thus, the mechanism of the self-organization in wedge diffraction is made clear by the VDPD analysis.

4:00

4pPA12. The phase gradient method (PGM) applied to a monoclinic plate immersed in water. Lionel Guénégou and Olivier Lenoir (LAUE UMR CNRS 6068, Univ. Le Havre, Pl. R. Schuman, 76610 Le Havre, France, lionel.guenegou@univ-lehavre.fr)

For a water-loaded isotropic plate, the PGM deals with the study of the partial derivatives of the phase of its reflection coefficient. The phase derivatives with respect to the frequency f, to the bulk phase velocities of the pressure (c_P) and shear (c_S) waves propagating in the plate, to the phase velocity c_F of the waves in water, and to the incident angle are investigated. The frequency and angular derivatives permit characterization of the frequency and angular resonances of the plate (locations and widths) without calculations in the associated complex planes. The derivatives with respect to $c_{P,S}$ give the prevailing polarization state of the Lamb waves. For a monoclinic plate (orthotropic plate where the incident plane is different from a symmetry plane), at a given incident angle, three phase velocities in the plate are to be considered. The additional one with regard to the isotropic case is the quasi-SH wave phase velocity. It is shown in this study that the frequency phase derivative study is still convenient to obtain the frequency resonance features, and that the phase derivatives with regard to the three velocities indicate whether a guided wave is mainly a quasilongitudinal, quasishear, or quasi-SH mode.

4:15

4pPA13. Flow noise and rapidly distorting turbulence in shear layers. R. Martinez (Cambridge Acoustical Assoc./Anteon Corp., 84 Sherman St., Cambridge, MA 02140)

Rapid distortion theory (RDT) is a linear analytical framework for the formal split of a flow's hydrodynamic sources from their acoustic effects. The former include time-varying vorticity, turbulence, and unsteady heat injection. Their acoustic effects manifest themselves through a generalized wave-propagation operator in the consistently separated equations of fluid mechanics. The present development reports on progress in extending RDT to include mean shear and associated macro-vorticity in the static but spatially nonuniform carrier flow. The analysis begins by fully recasting standard RDT [Goldstein, JFM (1978)] for irrotational backgrounds in tensor-dyadic form for the curvilinear coordinates of the streamlines of the background flow. This complete geometrization of the background offers clues for achieving a similar split in the time-varying acoustical and turbulent variables that perturb the more general sheared freestream. The new development could eventually be applied to turbulent boundary layers, and particularly to spatial discontinuities such as steps and gaps that (rapidly) distort the statistics of the perturbed rotational flow and thereby lead to additional broadband noise via RDT's generalized wave-equation operator and geometrized source terms.

Session 4pPP

Psychological and Physiological Acoustics and Speech Communication: The Perception of Complex Sounds: Honoring the Contributions of Charles S. Watson

Marjorie R. Leek, Cochair
Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, DC 20307-5001

Larry E. Humes, Cochair
Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405-7002

Invited Papers

1:00

4pPP1. "Watson, come here!" Larry Humes (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405-7002)

This famous quote marked a unique historical event in the development of telephony and an early experiment involving hearing. Although our own Watson postdates that work (considerably!), the life and work of Charles S. Watson are true reflections of good science and significant growth in our knowledge of human hearing. On behalf of those speaking at and attending this session in honor of Charles S. Watson, we ask this Watson to ''come here!'' (or, perhaps as a colleague suggested, to ''come hear!''). We ask that he do so to permit us to pay tribute to the many ways he has influenced the fields of psychoacoustics and speech communication, in general, and our careers, in particular. This introduction will provide a very brief overview of Chuck Watson, the person, as well as the scholar. By doing so, this overview will lay a foundation for understanding the impact of this man and his work. Subsequent speakers will build upon this foundation, with special focus on Chuck's work in the area of the perception of complex sounds.

1:15

4pPP2. The role of stimulus uncertainty in speech perception. Diane Kewley-Port (Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, kewley@indiana.edu)

Among the important experimental factors that affect psychophysical measurements of speech perception is stimulus uncertainty. Charles Watson has defined stimulus uncertainty as variation in stimulus parameters from trial to trial and demonstrated its highly degrading effects on a variety of complex auditory signals. Watson, Kelley, and Wroton showed large ($\times 10$) elevation of frequency-discrimination thresholds for ''word-length tonal patterns'' under high uncertainty conditions [J. Acoust. Soc. Am. **60**, 1176–1186 (1976)]. Investigations of speech, such as the perception of VOT (voice onset time) in stops [Kewley-Port, Watson, and Foyle, J. Acoust. Soc. Am. **83**, 1113–1145 (1988)] and discrimination of vowel formants [Kewley-Port, J. Acoust. Soc. Am. **110** (2001)], have also demonstrated the systematic and profound effects of higher levels of stimulus uncertainty. This presentation will discuss extensions of the concept of stimulus uncertainty that demonstrate the degrading effects of the variability in more natural speech (versus synthetic speech) and longer phonetic context (including sentences) on vowel formant discrimination. Results from normal-hearing and hearing-impaired listeners demonstrating similar detrimental effects of high stimulus uncertainty will also be presented. [Research supported by NIH-NIDCD.]

1:40

4pPP3. Molecular psychophysics and sound-source identification. Robert A. Lutfi (Dept. of Commun. Disord. and Waisman Ctr., Univ. of Wisconsin, Madison, WI 53706, ralutfi@wisc.edu)

Threshold and d-prime measures of performance are ubiquitous in psychophysics. Yet, because these measures require averaging over many responses they can conceal important aspects of the subject's decision process as it is reflected in the data from trial to trial. Chuck Watson clearly demonstrated this some 4 decades ago in his Ph.D. thesis. He was an early advocate of a ''molecular'' approach to psychophysics that attends specifically to the relation between individual stimuli and their associated response on each trial. In recent years, a variant of the molecular approach, perturbation analysis, has been applied with great success to the problem of image identification in vision [J. Vision **2**(1) (2002), special issue]. This talk reviews the application of this method to a similar longstanding problem in auditory psychophysics—the identification of sound sources. Published and previously unreported studies are presented that use synthesized sounds to investigate listener identification of the material, geometric, and driven properties of simple resonant sources—stretched membranes, clamped bars, and suspended plates. These studies show that when listeners are confronted with a complex identification task, for which there are multiple sources of acoustic information, they regularly adopt different decision strategies that yield the same level of identification accuracy (same d-prime). [Work supported by NIDCD.]

2:05

4pPP4. Neural mechanisms for spectral segregation. Donal G. Sinex (Dept. of Speech and Hearing Sci., Arizona State Univ., Box 870102, Tempe, AZ 85287-0102)

Humans exhibit a remarkable ability to segregate sounds produced by multiple sources that overlap in frequency and in time. It is likely that other species that use hearing also have this ability. To investigate the neural mechanisms that contribute to spectral segregation, responses of auditory nerve fibers (ANFs) and inferior colliculus (IC) neurons to harmonic complex tones and to the same tones with a mistuned component were measured. Mistuning leads to the perception of a new sound source and also produces dramatic qualitative changes in the temporal discharge patterns of IC neurons. In contrast, the same stimulus manipulation produces only modest quantitative changes in the responses of ANFs. These results indicate that the processing of complex tones undergoes a major transformation in the lower brainstem that is likely to contribute to perceptual segregation based on harmonicity. A computational model has been developed to investigate integrative mechanisms that may underlie this transformation. The model reproduces the distinctive discharge patterns of IC neurons, and suggests that these patterns arise as a result of narrow-band envelope extraction, followed by broadband excitatory–inhibitory interactions. Specific predictions of the model have been confirmed in subsequent electrophysiological experiments, providing further support for this interpretation. [Work supported by NIDCD.]

2:30

4pPP5. Chuck Watson's "differential psychoacoustics:" Individual differences in auditory abilities. Gary R. Kidd (Dept. of Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47404)

Chuck Watson was among the first in the psychoacoustic community to seriously address the topic of individual differences. At a time when there was little concern with variation among ''normal listeners'' in psychoacoustic research, Watson began a research program to document the range of human auditory abilities. The primary goals were to determine the number of distinct abilities, to specify the nature of each ability, and to document the distribution of these abilities in the general population. Thanks to Watson's talent for organizing and directing large-scale projects and his workmanlike approach to science, a large and valuable body of data on human individual differences has been collected. The research program began about 20 years ago with the study of basic auditory abilities, and it has expanded to include other modalities and cognitive/intellectual abilities in adults and children. A somewhat biased view of the importance of this work will be presented by one of Watson's many colleagues in this endeavor. The talk will provide an overview of this ongoing research program as well as a brief review of some related research by other investigators. New findings from recent extensions of this work will also be discussed.

2:55–3:10 Break

3:10

4pPP6. Understanding response proclivity and the limits of sensory capability: What do we hear and what can we hear? Marjorie R. Leek (Army Audiol. and Speech Ctr., Walter Reed Army Med. Ctr., Washington, DC 20307, leekmar@aol.com) and Donna L. Neff (Boys Town Natl. Res. Hospital, Omaha, NE 68131)

Charles Watson's studies of informational masking and the effects of stimulus uncertainty on auditory perception have had a profound impact on auditory research. His series of seminal studies in the mid-1970s on the detection and discrimination of target sounds in sequences of brief tones with uncertain properties addresses the fundamental problem of extracting target signals from background sounds. As conceptualized by Chuck and others, informational masking results from more central (even ''cogneetive'') processes as a consequence of stimulus uncertainty, and can be distinguished from ''energetic'' masking, which primarily arises from the auditory periphery. Informational masking techniques are now in common use to study the detection, discrimination, and recognition of complex sounds, the capacity of auditory memory and aspects of auditory selective attention, the often large effects of training to reduce detrimental effects of uncertainty, and the perceptual segregation of target sounds from irrelevant context sounds. This paper will present an overview of past and current research on informational masking, and show how Chuck's work has been expanded in several directions by other scientists to include the effects of informational masking on speech perception and on perception by listeners with hearing impairment. [Work supported by NIDCD.]

3:35

4pPP7. Learning about hearing by listening to Charles S. Watson for 45 years. James D. Miller (Dept. of Otolaryngol., Box 8115, Washington Univ. School of Medicine, 660 S. Euclid Ave., Saint Louis, MO 63110)

I met and started working with C. S. Watson on or about the Fall of 1958. Even then there was usually a pause and a wind-up before he made a point, usually neither concise nor terse. But once made, a point would be repeated and repeated. These oft-repeated points transform and become ''Watson's Precepts.'' Several such precepts, related to the study of speech perception by the hearing-impaired, will be quoted or paraphrased and their significance illustrated with samples from his work or that of others. Included on the list are the following. ''Bryan and Harter found that learning to perceive Morse Code continued over many hundreds of hours of practice. Training in many auditory tasks may require similar investments of practice time.'' ''One must distinguish response proclivities and sensory capabilities and tasks with low- or high-uncertainty stimulus conditions.'' ''Psychoacoustic measures, other than the audiogram, don't correlate with speech perception measures.'' Finally, it will be revealed how this collection of ''precepts'' has led to his current views of speech perception by the hearing impaired and to our current collaboration on auditory training for the hearing-aid users which features extensive practice and transition from low- to high-uncertainty stimulus conditions.

4:00

4pPP8. Studying environmental sounds the Watson way. Brian Gygi (East Bay Inst. for Res. and Education, 150 Muir Rd., Martinez, CA 94553)

After years of research on laboratory-generated complex sounds, in the early 1990s Chuck Watson and colleagues in the Hearing and Communications Laboratory (HCL) became interested in whether sounds with some meaning to the listener were processed differently by the auditory system. So began in his lab a program of environmental sounds research, in the meticulous, deliberate manner Watson was known for. The first step was developing an addition to the Test of Basic Auditory Capabilities (TBAC) which would measure individual differences in the identification of familiar environmental sounds. Next came the psychophysical basics: detection and identification in noise. Then, borrowing a page from early speech researchers, the effects of low-, high-, and bandpass filtering on environmental sounds were investigated, as well as those of processing environmental sounds using vocoder methods. Work has continued outside the HCL on developing a standardized canon of environmental sounds for generalized testing, with an aim to creating diagnostic tests for environmental sounds similar to the SPIN and modified rhyme and reverberation (MRRT).

4:15

4pPP9. Selective auditory attention to features of complex sounds: A comparative approach. Eduardo Mercado III (Dept. of Psych., Univ. at Buffalo, SUNY, Park Hall, Buffalo, NY 14260, emiii@buffalo.edu) and Itzel Orduña (Rutgers Univ., Newark, NJ 07102)

When listeners are trained to respond based on one spectrotemporal component of a complex sound, enhanced processing of the behaviorally relevant feature provides an objective correlate of selective attention [I. J. Hirsh and C. S. Watson, Annu. Rev. Psychol. **47**, 461–484 (1996)]. To study this issue in a nonhuman species, rats were trained to classify multidimensional acoustic stimuli based on the rate, direction, and range of frequency modulation. Rats successfully learned to classify complex sounds along the dimensions of rate and direction of frequency modulation, but not based on the range of frequency modulation. Rats classified stimuli most accurately when the relevant dimension was rate of frequency modulation. The relative ease with which rats learn to classify complex sounds along a particular dimension can be predicted based on how auditory cortical neurons in rats respond to such sounds. These findings provide new insights into how neural processing may constrain selective auditory attention to features of complex sounds. [Work supported by NIH.]

4:30

4pPP10. Listening weights for signals and maskers with uncertain frequency in normal-hearing and hearing-impaired listeners. Joshua M. Alexander and Robert A. Lutfi (Dept. of Commun. Disord. and Waisman Ctr., Univ. of Wisconsin, 1500 Highland Ave., Madison, WI 53705, jmalexa1@wisc.edu)

The effect of signal and masker uncertainty on listener decision weights was measured in 12 normal-hearing (NH) and 6 hearing-impaired (HI) listeners. The signal was a tone of 0.8, 2.0, or 5.0 kHz. The maskers were fixed-frequency tones separated from the signal(s) by two-third octaves and played simultaneously with the signal(s). Each masker tone had an independent probability of occurrence of 0.5 on each trial. In the signal-uncertain (SU) condition one of the three signals was played at random with equal probability on each trial. In the signal-certain (SC) condition the signal frequency was constant on each trial. In each condition the relative influence of each signal and masker frequency on listener decisions was estimated from regression weights relating listener responses to the presence or absence of each tone on every trial. For SC, the results indicate that NH and HI listeners put significantly greater weight on the signal frequencies than the masker frequencies. For SU, however, HI listeners tended to weight a narrower spectral region compared to NH listeners. It is concluded that HI listeners are able to selectively attend to an individual signal better than they can divide their attention among multiple signals. [Work supported by NIDCD.]

4:45

4pPP11. Contributions of internal noise and Bernoulli variance to the variability in multiple estimates of d'. Walt Jesteadt (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, jesteadt@boystown.org)

Multiple estimates of d' obtained from the same observer will vary as a result of differences in attention and other sources of internal noise, but also as a result of the variances associated with the two proportions that contribute to each d' estimate. This second source, known as Bernoulli variance, causes the expected variance of d' to vary a function of the true value of d'. Because estimates of d' obtained from 2×2 matrices are discrete rather than continuous, the expected variance of the estimates cannot be specified by an equation. Miller [Percept. Psychophys. **58**, 65 (1996)] has presented a method for computation of the sampling distribution of d', for any true value of d' and any given number of trials, and has demonstrated that a well-known approximation greatly exaggerates the variance for large values of d'. In the current paper, the standard approximation and Miller's exact method are extended from Yes–No to the more commonly used 2IFC procedure and the effects of two standard corrections for zero cells are examined. A comparison of the theoretical results to actual data suggests that Bernoulli variance plays a greater role than internal noise in determining the variability in d' estimates.

5:00

4pPP12. Informational masking of speech-analog signals: Independence of prosodic-rhythmic and formant trajectory cues. Pierre Divenyi (Speech and Hearing Res., VA Medical Ctr., Martinez, CA 94553, pdivenyi@ebire.org)

One factor of poor speech understanding in cocktail-party settings is masking of information in a target stream by similar concurrent information in an interfering background stream. To overcome this deficit, the information in the two streams must be different along at least some acoustic dimensions. The present study investigated whether informational differences along various dimensions between a speech-analog target and an interfering distractor stream were processed independently from one another, or whether dividing one's attention between the dimensions affects performance. In a single-interval, forced-choice experiment, listeners had to discriminate the rhythmic pattern of a burst triplet (AM) or an up–down/down–up 150-ms single-formant transition pattern (FM), or both, imposed on a 500-ms harmonic complex carrier, embedded in a distractor stream of random AM bursts, or constant single-formant up–down FM started in a random transition phase, or both. Fundamental frequencies of target and distractor were different. Results indicate that, when compared to discrimination of either the AM or the FM pattern alone, simultaneous discrimination of the AM and FM patterns did not result in a loss, suggesting that informational masking may operate independently on different dimensions. [Research supported by Grant R01-AG07998 from National Institute on Aging and by the VA Medical Research.]

Session 4pSA

Structural Acoustics and Vibration: Measurements and Transducers

Kenneth D. Frampton, Chair
Department of Mechanical Engineering, Vanderbilt University, Box 1592, Station B, Nashville, Tennessee 37235-1592

Contributed Papers

1:30

4pSA1. Noise analysis of a simplified Michelson interferometer vibrometer hydrophone. Lee E. Estes and Benjamin A. Cray (Naval Undersea Warfare Ctr., 1176 Howell St., Newport, RI 02841-1708)

The use of a Michelson interferometer to sense acoustic vibrations imposed on a water/plate/air boundary is investigated. To focus on fundamentals, a simplified interferometer is considered rather than the usual heterodyne or dual track homodyne configurations. Based on analytical models, the sensitivity limits due to photoelectron shot noise, laser light amplitude and phase noise, atmospheric turbulence, thermal vibration, and amplifier noise are predicted. Techniques for mitigation of turbulence and laser noise will be discussed.

1:45

4pSA2. Measurements of underwater acoustic pressure fields using a scanning laser Doppler vibrometer. Gerard P. Carroll (Naval Surface Warfare Ctr., Carderock Div., 9500 MacArthur Blvd., Bethesda, MD 20817-5700, carrollgp@nswccd.navy.mil)

Laser Doppler vibrometers (LDV) are designed to measure structural vibration velocity by sensing the phase shift in the laser signal reflected from a vibrating source. It is known that index of refraction modulations resulting from acoustic pressure distributions along a laser light path will also cause a phase shift. Simpson *et al.* [J. Acoust. Soc. Am. **99**(4), 2521(A) (1996)] have investigated this acousto-optic phase modulation as a possible contaminating effect for underwater LDV vibration measurements. This paper will investigate acousto-optic phase modulations measured by a scanning LDV as a method for measuring pressure radiating from underwater vibrating surfaces. This is done by passing the laser beam through the radiating pressure field and measuring the backscattered laser signal which is reflected off a rigid and retroreflective surface (outside the pressure field). It is shown experimentally, using the average pressure measured with an LDV over a plane in the vicinity of a vibrating structure, that the pressure at a far-field location normal to the plane can be determined.

2:00

4pSA3. Visualization of the energy flow for guided forward and backward waves in and around a fluid-loaded elastic cylindrical shell via the Poynting vector field. Cleon E. Dean (Phys. Dept., Georgia Southern Univ., P.O. Box 8031, Statesboro, GA 30460-8031, cdean@GeorgiaSouthern.edu) and James P. Braselton (Georgia Southern Univ., Statesboro, GA 30460-8093)

Color-coded and vector-arrow grid representations of the Poynting vector field are used to show the energy flow in and around a fluid-loaded elastic cylindrical shell for both forward- and backward-propagating waves. The present work uses a method adapted from a simpler technique due to Kaduchak and Marston [G. 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**, 3501–3507 (1995)] to isolate unidirectional energy flows.

2:15

4pSA4. Cylindrical transducers to generate and detect axisymmetric waves in a pipe. Jin O. Kim, Kyo-Kwang Hwang, Jung-Goo Lee, and Hyung-Gon Jeong (Soongsil Univ., 1 Sangdo-dong, Dongjak-gu, Seoul 156-743, Korea, jokim@ssu.ac.kr)

This paper presents the radial vibration characteristics of piezoelectric cylindrical transducers and the application of the transducers to the generation and detection of axisymmetric longitudinal waves in a pipe. Dynamic differential equations of piezoelectric radial motion derived in terms of radial displacement and electric potential and mechanical and electric boundary conditions have yielded a characteristic equation for radial vibration of the radially polarized piezoelectric cylinder. Theoretical calculations of the fundamental natural frequency have been compared with numerical and experimental results for transducers of several sizes, and have shown a good agreement. It has been shown that the piezoelectric natural frequency of the fundamental mode for radial vibration in a cylindrical transducer depends mostly on the radius rather than on the thickness of the cylinder. Experiments have been performed in empty and water-filled pipes equipped with the transducers that were used for transmitting and receiving axisymmetric elastic waves in the pipe wall. The measured wave speeds have been compared with the analytical ones. This work has demonstrated the feasibility of using cylindrical transducers and pipe waves for the determination of the mass density and, eventually, the flow rate of the liquid in a pipe.

2:30

4pSA5. Actuator/sensor placement optimization for vibration control based on finite element model of the car chassis. Bouzid Seba, Nikola Nedeljkovic, and Boris Lohmann (Inst. of Automation, Bremen Univ., Bremen 28359, Germany)

The finite element method is a convenient approach to model the structures with complex geometry and subjected to complicated boundary conditions. However, the order of the obtained model using FEM strategy is large. Therefore, the control design and implementation based on the FEM model turns out to be complicated (sometimes not feasible). Hence, the order reduction techniques are inevitable to reduce the size of the system obtained with FEM. The reduced model is then used to explore and to optimize the system in terms of collocated actuator/sensor positioning problem. In this paper, we suggest a strategy to optimize the location of the electromechanical actuators in the car chassis subframe for vibration control. The proposed method can be formulated using an optimization problem with constraints which are introduced to consider the spill over effect resulted from the neglected higher modes in the model reduction part (which affects the stability of the closed loop system) and to ensure the minimum controllability of individual modes. The optimization algorithm is used to select three optimal positions of the collocated actuator/sensor among 298 possible positions.

2:45

4pSA6. The influence of the piezoelectric actuators distribution on the plate acoustic radiation. Jerzy W. Wiciak (Structural Acoust. and Biomed. Eng. Lab., Univ. of Sci. and Technol., Poland)

The analysis of sound radiation with steady-state vibration of a steel plate with the system of piezoceramic elements was performed.The system of piezoceramic actuators consists of four pairs PZT4 squared elements with dimensions $10\times10\times1$ mm^3. The actuators were located symmetrically on the plate and driven 180 deg out of phase with the same signal. It was assumed that rectangular plate is fixed along one edge. Other edges have the free boundary conditions case. The structural vibration analysis of a plate with a system of piezoelectric actuators was performed with FEM ANSYS@ code for the first three modes. The hybrid [Kozien and Wiciak, Quantum Mol. Acoust. **24**, 98–110 (2003)] and fluid structure interactive methods were applied to estimate the sound pressure radiated by a vibrating system plate and actuators. Analysis of the SPL radiated by a vibrating plate with piezoceramic actuators into chosen points in the acoustic field gave similar results with two applied methods. The small differences in results were possibly an effect of the boundary assumptions. The best results in the reduction of sound radiated by plate were obtained for the first mode. [Work supported by SCfFR.]

3:00–3:15 Break

3:15

4pSA7. Error analysis and implementation issues for energy density probe. Lance L. Locey, Brady L. Woolford, Scott D. Sommerfeldt, and Jonathan D. Blotter (N283 Eyring Sci. Ctr. (ESC), Provo, UT 84602)

Previous research has demonstrated the utility of acoustic energy density measurements as a means to gain a greater understanding of acoustic fields. Three spherical energy density probe designs are under development. The first probe design has three orthogonal pairs of surface mounted microphones. The second probe design utilizes a similarly sized sphere with four surface mounted microphones. The four microphones are located at the origin and unit vectors of a Cartesian coordinate system, where the origin and the tips of the three unit vectors all lie on the surface of the sphere. The third probe design consists of a similarly sized sphere, again with four surface microphones, each placed at the vertices of a regular tetrahedron. The sensing elements of all three probes are Panasonic electret microphones. The work presented here will expand on previously reported work, and address bias errors, spherical scattering effects, and practical implementation issues. [Work supported by NASA.]

3:30

4pSA8. Free field reconstruction using energy density measurements. Michael C. Harris, Jonathan D. Blotter, and Scott D. Sommerfeldt (Dept. of Mech. Eng., Brigham Young Univ., 435 CTB, Provo, UT 84602)

Forward wave propagation to the farfield using nearfield pressure measurements has undergone extensive development. Up until now, discrete pressure measurements are typically made using a conformal array of microphones. However, complex acoustic fields require large microphone arrays in order to accomplish accurate field reconstruction. The use of energy density sensors has indicated several advantages over conventional pressure microphones in active noise control applications. This study examines possible benefits of energy-based measurements in free field reconstruction. To determine and quantify the possible advantages of energy density measurements, an analytical model was built to evaluate the performance. The error in acoustic fields reconstructed using energy density sensors is compared with conventional pressure microphone reconstruction accuracy. A number of analytical studies will be presented showing the benefits of using energy-based measurements in free field reconstruction.

3:45

4pSA9. Detecting and classifying adhesive flaws between bonded elastic plates. Ricardo Leiderman (Dept. Engenharia Mecânica, Pontificia Universidade Catolica do Rio de Janeiro, PUC-Rio, Brazil) and Paul E. Barbone (Boston Univ., Boston, MA 02215)

Nondestructively evaluting the quality of an adhesive bond is challenging. The reason is that the only way to measure bond strength is by measuring the force required to break a bond, after which the bond is broken. Localized adhesion flaws that diminish bond strength, however, also tend to diminish bond stiffness. For example, pockets of elevated porosity, microcracking, or other damage in the adhesive layer will simultaneously lower local bond strength and local effective bond stiffness. So guided, we formulate and solve an inverse scattering problem to reconstruct the effective stiffness *distribution* of an adhesive layer in a layered elastic plate. Our formulation is based on the method of invariant imbedding, and applies to isotropic and anisotropic elastic layers. We present two solutions of the inverse problem: the Born approximation and the exact solution. Both solutions are unstable, as is the nature of inverse scattering problems, and require some regularization in the presence of noise. We present computed examples and discuss the role of regularization.

4:00

4pSA10. Assessment of material fatigue damage using nonlinear vibro-modulation technique. Andrei Zagrai (Davidson Lab., Stevens Inst. of Technol., 711 Hudson St., Hoboken, NJ 07030, azagrai@stevens.edu), Dimitri Donskoy (Intelligent Sensing Technologies, LLC, Fair Haven, NJ 07704-6457), Alexander Chudnovsky, and Hudson Wu (Univ. of Illinois at Chicago, Chicago, IL 60607)

Heavy periodic loads exerted on structural materials often lead to fatigue damage (material degradation at microscale) which may finally trigger irreversible fracture process. Conventional NDT techniques detect only the latter, and there is an increasing need for new tools to assess fatigue damage at the earliest possible stage, i.e., before fracture. This paper presents experimental results of early damage characterization using an innovative nonlinear vibro-modulation technique (VMT) [Donskoy *et al.*, NDT&E Int. **34** (2001)]. In the experiments, fatigue damage was initiated in steel, aluminum, and carbon–carbon composite specimens during strain-controlled three-point bending high-cycling fatigue tests. The damage progress was independently monitored using dataflow from the testing machine and the real-time nonlinear vibro-modulation measurements. The tests demonstrated that the reduction in the specimens' stiffness (direct indication of damage accumulation) correlates well with the increase in the VMT's nonlinear damage index. These results confirm that VMT could offer new opportunities for early damage detection and remaining life prediction. [Work supported by NAVAIR.]

4:15

4pSA11. The application of acoustical methods of ply separation detection in plates. Marek Iwaniec (Univ. of Sci. and Technol., Structural Acoust. and Biomed. Eng. Lab., Al. Mickiewicza 30, 30-059 Krakow, Poland)

The author applied modal analysis and energy flow methods for detection of multilayer plate delaminating. New design of hardware for field measurements especially for conservation of monuments of history was developed. The author also tested two new noninvasive measurement methods. The first one is based on measurement of electric field dynamic parameters in the vicinity of subsurface delaminating and the second on the measurement of the frequency response of a structure being excited by the diffusive acoustic near field of the radiator. Main application of developed devices was use in a wide band of practical applications in detection of delaminating of precious wall paintings—frescoes. Preliminary tests at laboratory conditions were carried out with designed devices on fresco samples prepared due to old technologies. Results proved satisfactory sen-

sitivity of the devices for practical application to conservator team works. The assessment of a structure technical state is performed on the basis of real-time calculations of the quality measures and damage probability. The proposed device enables the investigator to identify fresco delaminating of 8-cm diameter located at a depth of 5 mm. Discussed methods can be useful in diagnostics of many other multilayer structures.

THURSDAY AFTERNOON, 27 MAY 2004 CONFERENCE ROOM E, 1:40 TO 3:30 P.M.

Session 4pSP

Signal Processing in Acoustics: Sensor and Array Processing

Paul Hursky, Chair
SAIC Ocean Sciences Division, 10260 Campus Point Drive, San Diego, California 92121

Chair's Introduction—1:40

Contributed Papers

1:45

4pSP1. Model-based broadband towed-array processing. Edmund J. Sullivan (OASIS, Inc., 5 Militia Dr., Lexington, MA 02421) and Geoffrey S. Edelson (BAE SYSTEMS, Nashua, NH 03061-0868)

Conventional broadband towed array processors operate in the frequency domain and exploit the phases of the frequency components from a DFT to extract the bearing estimate. In this process, the fact that the array is moving is not taken into consideration. In fact, to first order, the effect of the motion is negligible. However, it has been shown that, in the narrow-band case, the effect of the motion on the variance of the bearing estimate is not negligible. This is due to the fact that the Doppler incurred by the motion itself carries bearing information. In this work, it is shown that a recursive bearing estimation scheme, based on sequential DFT's, can produce the so-called synthetic aperture effect, where the Cramer–Rao lower bound on the bearing estimate progressively decreases, as compared to the case where the motion is not explicitly included in the processing scheme. Examples of this effect are shown based on simulated data, and the capabilities and limitations of the technique are discussed.

2:00

4pSP2. Optimal sensor placement in highly variable noise fields to increase detection. Donald R. DelBalzo and Erik R. Rike (Neptune Sci., Inc., 40201 Hwy. 190 E., Slidell, LA 70461, delbalzo@neptunesci.com)

Ambient noise (AN) is highly variable in littoral environments. AN predictions are often smoothed to produce mean estimates in support of signal detection. In a complex AN environment, the optimal placement of sensors for detection also depends on the spatial and temporal variance of AN. The recent development of the Dynamic Ambient Noise Model (DANM) allows accurate, detailed estimation of noise statistics. DANM was used to produce space- and time series of directional noise based on discrete ship tracks. This work describes a fast optimization scheme that exploits nonhomogeneous noise variability to select optimal locations that maximize detection. A standard, nonoptimal approach for choosing locations is sequential and myopic. A better approach is combinatorial; i.e., consider all possible sensor sets and choose the best. This optimal approach is often computationally prohibitive. Brown's iterative algorithm combines elements of both algorithm types to achieve rapid results. Brown's algorithm was modified and used to demonstrate the importance of nonhomogeneous noise statistics. The results show a significant improvement in probability of detection when DANM AN statistics are considered while choosing optimal sensor locations. This work has application to design of optimal sonobuoy field patterns. [Work sponsored by ONR under the LADC project.]

2:15

4pSP3. Blind deconvolution for non-stationary acoustical systems. Mark R. Gramann, Josh G. Erling, and Michael J. Roan (Appl. Res. Lab, Penn State Univ., P.O. Box 30, State College, PA 16804, mrg227@psu.edu)

Blind deconvolution is a signal-processing technique that has been shown to be highly effective for removing multipath propagation channel-induced corruption of a signal. These algorithms are usually implemented as adaptive filtering processes that learn finite impulse response inverse filters based on measurements of the corrupted signal. A major limitation of these techniques is that a stationary propagation environment is generally assumed. This assumption is not valid for many real-world applications, such as active or passive sonar operating in a shallow-water environment or for communications systems with moving sources or receivers. For slowly varying systems, standard algorithms with adaptive learning rates that improve convergence speed, such as frequency-domain implementations of the Infomax and Natural Gradient algorithms, may be sufficient. However, in systems where either the input signal or impulse response change rapidly, it becomes necessary to provide a means for tracking the statistical changes that occur in the input signal or impulse response. Approaches to tracking these changes through the use of a tracking algorithm such as Kalman filtering are discussed. [Work supported by Dr. David Drumheller, ONR Code 333, Grant No. N00014-00-G-0058.]

2:30

4pSP4. Free-field equivalent shape of scattering conformal microphone arrays. Philippe Moquin and Stephane Dedieu (Mitel Networks, 350 Legget Dr., Kanata, ON K2K 1X3, Canada)

When an array is placed upon an acoustic scatterer, the interelement spacing of the array appears to be different from free field conditions [J. Meyer, J. Acoust. Soc. Am. **109**, 185–193 (2001)]. This manifests itself primarily as a nonlinearity in the interelement phase difference. Ideal scatterers (rigid and nonconducting) can be used to understand these effects. A spherical scatterer with a circular array serves as a first example since it has an analytical solution. To help understand the effect, equivalent shape of a free-field array can be calculated. Numerical simulation and measured results on a telephonelike shape with an elliptical array will serve to illustrate that the results from the sphere can be used as a guide for estimating the scattering effects that can be expected in a more complex shape.

2:45

4pSP5. A decentralized algorithm for acoustic localization using a distributed sensor network. Peter L. Schmidt, Isaac Amundson, and Kenneth D. Frampton (Dept. of Mech. Eng., Vanderbilt Univ., VU-Station B, Nashville, TN 37212)

An acoustic source localization algorithm has been developed for use with large-scale sensor networks using a decentralized computing approach. This algorithm, based on a time delay of arrival (TDOA) method, uses information from the minimum number of sensors necessary for an exactly determined solution. Since the algorithm is designed to run on computational devices with limited memory and speed, the complexity of the computations has been intentionally limited. The sensor network consists of an array of battery-operated COTS Ethernet-ready embedded systems with an integrated microphone as a sensor. All solutions are calculated as distinct values, and the same TDOA method used for solution is applied for ranking the accuracy of an individual solution. Repeated for all combinations of sensor nodes, solutions with accuracy equivalent to complex array calculations are obtainable. Effects of sensor placement uncertainty and multipath propagation are quantified and analyzed, and a comparison to results obtained in the field with a large array with a centralized computing capability using a complex, memory intensive algorithm is included.

3:00

4pSP6. Noise reduction using a multimicrophone array for automatic speech recognition on a handheld computer. Scott T. Shaw, Andrew J. LaRow, William E. Schoenborn, Jason Rodriguez, and Gary L. Gibian (Planning Systems, Inc., 12030 Sunrise Valley Dr., Reston, VA 20191-3453, ggibian@plansys.com)

A four-microphone array and signal-processing card have been integrated with a handheld computer such that the integrated device can be carried in and operated with one hand. Automatic speech recognition (ASR) was added to the USAMRMC/TATRCs Battlefield Medical Information System (BMIST) software using an approach that does not require modifying the original code, to produce a Speech-Capable Personal Digital Assistant (SCPDA). Noise reduction was added to allow operation in noisier environments, using the previously reported Hybrid Adaptive Beamformer (HAB) algorithm. Tests demonstrated benefits of the array over the HP/COMPAQ-iPAQ built-in shielded microphone for noise reduction and automatic speech recognition. In electroacoustic and human testing including voice control and voice annotation, the array provided substantial benefit over the built-in microphone. The benefit varied from about 5 dB (worst-case scenario, diffuse noise) to about 20 dB (best-case scenario, directional noise). Future work is expected to produce more rugged SCPDA prototypes for user evaluations, revise the design based on user feedback and real-world testing, and possibly to allow hands-free use by using ASR to replace the push-to-talk switch, providing feedback aurally and/or via a head-up display. [Work supported by the U.S. Army Medical Research and Materiel Command (USAMRMC), Contract No. DAMD17-02-C-0112.]

3:15

4pSP7. Plane-wave decomposition by spherical-convolution microphone array. Boaz Rafaely (Elec. and Computer Eng. Dept., Ben-Gurion Univ., Beer-Sheva 84105, Israel, br@ee.bgu.ac.il) and Munhum Park (Univ. of Southampton, Southampton SO17 1BJ, UK)

Reverberant sound fields are widely studied, as they have a significant influence on the acoustic performance of enclosures in a variety of applications. For example, the intelligibility of speech in lecture rooms, the quality of music in auditoria, the noise level in offices, and the production of 3D sound in living rooms are all affected by the enclosed sound field. These sound fields are typically studied through frequency response measurements or statistical measures such as reverberation time, which do not provide detailed spatial information. The aim of the work presented in this seminar is the detailed analysis of reverberant sound fields. A measurement and analysis system based on acoustic theory and signal processing, designed around a spherical microphone array, is presented. Detailed analysis is achieved by decomposition of the sound field into waves, using spherical Fourier transform and spherical convolution. The presentation will include theoretical review, simulation studies, and initial experimental results.

THURSDAY AFTERNOON, 27 MAY 2004 NEW YORK BALLROOM B, 1:25 TO 5:00 P.M.

Session 4pUW

Underwater Acoustics: Acoustic Propagation and Modeling

Michael Brown, Chair
RSMAS-AMP, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149-1098

Chair's Introduction—1:25

Contributed Papers

1:30

4pUW1. Focusing at an arbitrary waveguide location using time reversal. Shane C. Walker, Philippe Roux, and W. A. Kuperman (Marine Physical Lab., Scripp's Inst. of Oceanogr., UCSD, 8820 Shellback Way, La Jolla, CA 92093, shane@physics.ucsd.edu)

A method for producing acoustic time reversal focusing at an arbitrary location is presented. In principle, producing a focus at any location in a waveguide is trivial once the point to point propagation matrix is known. In practice, the propagation matrix over a given range, $G(zs,zr,R)$ between sources zs at $r=0$ and receivers zr at $r=R$, can be directly measured with a pair of vertical line arrays. Once known, repeated iterations of $G(zs,zr,R)$ result in good approximations to the propagation matrices, $G(zs,zr,mR)$, for ranges mR, where $m=2,3,\ldots$ is the number of iterations applied. By combining this iteration technique with frequency-dependent range shifting, it is possible to produce acoustic time reversal focusing at any range and depth.

1:45

4pUW2. Diffraction of nonsinusoidal sound in the sea. John L. Spiesberger (Dept. of Earth and Environ. Sci., 240 S. 33rd St., Univ. of Pennsylvania, Philadelphia, PA 19104, johnsr@sas.upenn.edu)

A theory of diffraction is applied to nonsinusoidal emissions of low-frequency sound in the sea. The ray approximation is sometimes good and other times not so good. Some theories for the scattering of nonsinusoidal emission of sound assume that propagation occurs within a Fresnel zone

of the ray path. This assumption is incorrect. Indeed, the use of a correctly sized region based on a theory of diffraction appears to be able to help resolve major discrepancies between observations and such theories for acoustic fluctuations. Indeed, the correctly sized region can be orders of magnitude smaller than given by a Fresnel radius.

2:00

4pUW3. Ray and travel time stability in weakly range-dependent sound channels. Francisco J. Beron-Vera and Michael G. Brown (RSMAS-AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

Ray path and travel time stability are investigated in environments consisting of a range-independent background on which a weak range-dependent perturbation is superimposed. Theoretical arguments suggest and numerical results confirm that both ray path and travel time stability are strongly influenced by the background sound speed profile. Both ray path and travel time stability are shown to increase with increasing magnitude of $\alpha(I)=(I/\omega)d\omega/dI$, where $2\pi/\omega(I)$ is the range of a ray cycle (double loop in a deep ocean sound channel) and I is the ray action variable. This behavior is illustrated using internal-wave-induced scattering in deep ocean environments and rough surface scattering in upward refracting environments. [Work supported by ONR.]

2:15

4pUW4. Wavefield stability in weakly range-dependent sound channels. Michael G. Brown, Francisco J. Beron-Vera, Irina Rypina, and Ilya Udovydchenkov (RSMAS-AMP, Univ. of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149)

It is shown that the mode-based "waveguide invariant" β is asymptotically equivalent to the ray-based "stability parameter" α. These parameters are known to control the dispersive properties of the sound channel and various measures of both ray path and travel time stability. This leads to the expectation that finite-frequency wavefield stability is also largely controlled by β (or α). PE simulations are shown to confirm this expectation. [Work supported by ONR.]

2:30

4pUW5. Three-dimensional wide-angle azimuthal PE solution to modified benchmark problems. Li-Wen Hsieh, Ying-Tsong Lin, and Chi-Fang Chen (Dept. of Eng. Sci. and Ocean Eng., Natl. Taiwan Univ., No. 73 Chou-Shan Rd., Taipei, Taiwan 106, R.O.C., rhs@ms11.url.com.tw)

In predicting wave propagation, the size of the angle of propagation plays an important role; thus, the concept of wide angle is introduced. Most existing acoustic propagation prediction models do have the capability of treating the wide angle but the treatment, in practice, is vertical propagation angle. This is desirable for solving 2D $(r-z)$ problems. Typically, 3D problems are dealt with an N by 2D approximation. To deal with problems possessing 3D effects, the azimuthal coupling terms have to be considered in PE approximation. Moreover, in extending the 2D treatment to 3D, the wide−angle capability is maintained in most 3D models, but it is still vertical. Hence, the concept of wide angle is introduced in the azimuthal direction to enhance the capability of predicting the azimuthal coupling and thus the 3D effects. A truncated wedge−shaped ocean which is modified from ASA benchmark problems is used in this study. The results show apparent 3D effects and validate the 3D wide−angle azimuthal PE model, a wide−angle version of FOR3D.

2:45

4pUW6. Approaches to accelerate three-dimensional acoustic propagation models. Juan Arvelo, Jr. and Allan Rosenberg (Appl. Phys. Lab., Johns Hopkins Univ., Laurel, MD)

A three-dimensional split-step Pade parabolic-equation approach [J. Comput. Acoust. **9**(2), 17–39 (2001)] was developed as an extension to Collins and Chin-Bings algorithm [J. Acoust. Soc. Am. **87**, 1104–1109 (1990)] to account for azimuthal acoustic refraction via a narrow-angle approximation. It was shown that high accuracy requires fine azimuthal and range step sizes. Therefore, a couple of Pade terms are sufficient given the required small range step size. The paper also mentions the possible implementation of a variable azimuthal grid size, where the number of azimuthal segments increases proportionally with increasing range, as another approach to boost the models computational speed. In this talk, we investigate the possibility of azimuthal padding as another approach to speed-up the model when the user is only interested in propagating the acoustic field along a wedge of azimuthal angles instead of the entire 360.

3:00

4pUW7. A variable rotated parabolic equation for elastic media. Donald A. Outing, William L. Siegmann (Rensselaer Polytechnic Inst., Troy, NY 12180, outind@rpi.edu), and Michael D. Collins (Naval Res. Lab., Washington, DC 20375)

There are several approaches for improving the accuracy of the parabolic equation method for problems involving sloping interfaces. Energy-conservation and single-scattering corrections have essentially resolved the acoustic problem, but these approaches have been less effective for problems involving elastic layers. Some promising results have recently been obtained using approaches based on mapped [J. Acoust. Soc. Am. **107**, 1937–1942 (2000)] and rotated [J. Acoust. Soc. Am. **87**, 1035–1037 (1990)] coordinates. The latter of these approaches provides greater accuracy for problems involving relatively steep interfaces. With the extension to variable slopes [J. Acoust. Soc. Am. **114**, 2428–2429 (2003)], the rotated parabolic equation became a useful approach for solving a large class of range-dependent problems. This presentation will describe a generalization of the variable rotated parabolic equation to elastic media. At each change in slope, a new rotated coordinate system is introduced and the solution is interpolated onto the new grid. In the elastic case, this process requires special operators for changing variables, the rotation of the dependent variables, and a rotationally invariant quantity. Examples will be presented to illustrate and test the approach. [Work supported by the ONR.]

3:15–3:30 Break

3:30

4pUW8. Broadband normal-mode computations within a multiprocessing environment. Steven A. Stotts and Field G. Van Zee (Appl. Res. Labs., Univ. of Texas, 10000 Burnet Rd., Austin, TX 78758, stotts@arlut.utexas.edu)

A technique to facilitate normal-mode modeling within a multiprocessing environment is presented using the Message-Passing Interface (MPI) communication library. Portability and standardization were the driving factors for using MPI as the multiprocessor implementation. Using only a minimal set of commands, an example algorithm illustrates the approach as a basis to transform existing models to perform computations via multiprocessors. Applications to propagation modeling are given with an emphasis on broadband normal-mode computations. A static approach to equalizing processor workload for normal-mode computations utilizes the linear dependence of mode number with frequency. The area under a curve representing the mode number versus frequency is used to determine the loading. An assumption of equal area implying equal loading determines the beginning and ending frequency increments for each processor. Comparisons with other approaches to processor loading for optimizing speedup time are given. A determination of the optimal number of processors for several examples of broadband normal-mode computations using a subroutine version of the Navy Standard NAUTILUS will be presented. Implementing this approach into more advanced models such as geoacoustic inversions will be discussed.

3:45

4pUW9. Progress in hybrid finite-element/propagation tool modeling of scattering from objects in underwater waveguides. Mario Zampolli, David S. Burnett, Finn B. Jensen, Alessandra Tesei (NATO Undersea Res. Ctr., Viale San Bartolomeo 400, 19138 La Spezia, Italy, zampolli@saclantc.nato.int), Henrik Schmidt (MIT, Cambridge, MA 02139), and John B. BlottmanIII (Naval Undersea Warfare Ctr., Newport, RI 02841)

A steady-state 3D finite-element software called FESTA (Finite-Element STructural Acoustics), is being developed at the NATO Undersea Research Centre. The code is geared towards a variety of applications in underwater acoustics, such as multistatic scattering from localized inhomogeneities, scattering across interfaces between fluids and/or solids, and multistatic scattering from single and multiple fluid-loaded elastic targets. One issue of importance to researchers in underwater acoustics is the trade-off between full 3D models and thin-shell theory. To address this issue, FESTA results are compared to results from thin-shell scattering codes for cylindrical bodies with spherical endcaps. Another part of the work focuses on the integration of the finite-element tool with underwater propagation tools for the computation of multistatic scattering from targets buried, partially buried or proud inside shallow-water waveguides. To achieve this, FESTA is coupled to the MIT underwater propagation tool OASES. One method of interfacing FESTA to OASES is based on an environment-independent characterization of the target via *in vacuo* response matrices. Recent results from response matrix computations are presented and trade-offs of this method are discussed.

4:00

4pUW10. Multiple uncorrelated scattering effects in waveguide propagation. Brian H. Tracey (MIT Lincoln Lab., 244 Wood St., Lexington, MA 02420, btracey@ll.mit.edu)

First-order perturbation theory has been widely used to model scattering from surface roughness in ocean waveguides. After sound is initially scattered it will propagate further and may rescatter from the rough surface. This rescattering effect is not captured by first-order perturbation theory. Voronovich [*Wave Scattering from Rough Surfaces* (Springer-Verlag, New York, 1994)] has argued that the scattered energy will return to the rough interface at range intervals that are large compared to the correlation length of the roughness. Thus the different scattering events can be treated as uncorrelated and a single-scattering theory can be used to describe each event. In this paper an existing normal mode scattering theory is modified to include this extension to uncorrelated rescattering. Numerical examples for low frequencies in shallow water are used to understand how the importance of the rescattering depends on seabed type. Rescattering was not found be significant for sandy bottoms, but is predicted to become more important as the bottom compressional speed increases and more scattered energy is trapped in the waveguide. [Work sponsored by DARPA-ATO under Air Force Contract F19628-00-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.]

4:15

4pUW11. Benchmark comparisons of a modified coupled normal mode method (Swamp) and commonly used range dependent methods. Michael Werby (NRL Code 7181, Stennis Space Ctr., MS 39529)

The algorithm Swamp has been benchmarked and documented and has been effectively used in a variety of studies [see http://www.ucs.louisiana.edu/~nxs7560]. Here the method is augmented in two ways. First, a more general layering scheme is implemented which allows for a broad angular spectrum required to replicate near field effects. Second, by making use of pseudo unitary operators the method is extended to a two-way coupled mode algorithm. The method is tested and compared with available methods in current use. Further, it is possible to make use of bandedness and this improves the time of calculation significantly and allows the resulting code to be used for very high frequencies. Benchmark calculations are presented for several examples.

4:30

4pUW12. Statistical modeling of sonar performance and its variance due to ocean fluctuations. Chi-Fang Chen, Zheng-Ru Yang, Mei-Chun Yuan, and Yuan-Ying Chang (Dept. of ESOE, Natl. Taiwan Univ., No. 1 Roosevelt Rd., Section 4, Taipei, Taiwan 106, R.O.C., chifang@ntu.edu.tw)

Sonar performance can be characterized by its detection range, which is determined by the sonar equation. The sonar equation includes parameters of sonar systems, environments, and the detection algorithms. Though the detection algorithm is itself a statistical means, it does not describe the relationship between the statistical properties of the environmental parameters with the detection range. This paper presents a statistical model for detection range based on the statistical properties of the environments, namely transmission losses. Temperature data from local moorings and VLA in ASIAEX-SCS are used to establish an ensemble samples for sound-speed profiles, and transmission loss curves are generated, thus the ensemble samples for TLs. The resultant detection range and its error bar are estimated for the Continental Shelf region in the South China Sea.

4:45

4pUW13. Variation of the acoustic directivity pattern due to inhomogeneous objects in the region of nonlinear interaction. Nikolai Zagrai (Taganrog State Univ. of Radio-Eng., GSP-17A, 44 Nekrasovsky St., Taganrog 347928, Russia, znp@tsure.ru)

The homogeneity of the physical medium can be investigated using nonlinear interaction of the acoustic waves. The proposed approach utilizes the fact that the secondary acoustic field is dependent on configuration of the region of nonlinear interaction (RNI), inhomogeneity of its physical properties and the transmitting signals of the primary acoustic field. Therefore, by studying a directivity pattern of the secondary acoustic field at combination frequencies it is possible to determine a number and location of the inhomogeneous objects in the RNI. The axial position and transverse dimensions of the objects affect the directivity pattern measured in the far field. In the case of the axially symmetric RNI, displacement of an object leads to the asymmetry of the directivity pattern and can be easily estimated. The model is proposed to account for the respective effect of the objects size, degree of asymmetry, and dynamics of the directivity pattern at the combination frequencies. The developed approach suggests that the methods of nonlinear acoustics can be effectively used for physical diagnostics and condition monitoring.

Session 5aAA

Architectural Acoustics: Topics in Architectural Acoustics

Benjamin C. Mueller, Chair
Ostergaard Acoustical Associates, 200 Executive Drive, West Orange, New Jersey 07052

Chair's Introduction—9:00

Contributed Papers

9:05

5aAA1. Investigation of early reflections on the perception of audio. John Gorr (Columbia College Chicago, 1025 W. Wood St., Palatine, IL 600067, jgorr1@msn.com)

The effects of early reflections in a nonsymmetrical control room were measured using time delay spectrometry and MLS methods. The tests revealed that comb-filtering patterns were the dominant source of spectral distortion. The mixing console reflections were investigated to reveal the subjective impact of the comb-filtering at the mix position. Tests were performed with and without acoustic foam on the console to determine the influence of the early reflections and then compared with theoretical values for the filtering patterns. Results were compared and used to qualify the subjective perception of the distortion.

9:20

5aAA2. Sound absorption and reflection from ceilings in open offices. Alf Warnock (Natl. Res. Council Canada, M59, 1200 Montreal Rd., Ottawa, ON K1A 0R6, Canada, alf.warnock@nrc-cnrc.gc.ca)

Sound attenuation between work areas in open offices depends on barrier height and the sound-absorbing properties of the ceiling. Two methods for evaluating the sound-absorption ceilings are available: ASTM E1111, which gives a rating called the articulation class (AC), and ASTM C423, which gives sound-absorption average (SAA). Measurements of six ceiling systems made in accordance with ASTM E1111 and ASTM C423 provide a link between AC and SAA. Manufacturers often do not provide both ratings for their products, and it is useful to be able to use either rating system. SAA and measured AC for a 1.5-m-high screen are related by $AC = 102 \times SAA + 91.4$. SAA and calculated values of AC for a 1.8-m-high screen are related by $AC = 118\ SAA + 93$. In the absence of E1111 test data, C423 SAA values can be used to select ceiling systems. The commonly recommended minimum SAA of 0.9 corresponds to an AC of 180 when the screen height is 1.5 m and to an AC of 200 when the screen height is 1.8 m. An approximate conversion from absorption coefficients to reflection coefficients agrees well with earlier work.

9:35

5aAA3. Prediction of broadband nonuniform time-dependent acoustic fields in enclosures with diffuse reflection boundaries using energy-intensity modes. Donald B. Bliss and Linda P. Franzoni (Mech. Eng. and Mater. Sci., Duke Univ., Durham, NC 27708, dbb@duke.edu)

A new analysis of high-frequency broadband reverberant sound fields in rooms with diffuse reflection boundaries is described. Depending on shape, source location, and distribution of wall absorption, rooms exhibit spatial variation in steady-state mean-square pressure and also spatial dependence of decay time characteristics. The room boundaries can be replaced by a distribution of uncorrelated broadband directional energy-intensity sources. In steady state with diffuse reflection boundaries, the interior pressure field produced by these sources satisfies Laplace's equation. The mean-square pressure field is expressed as a sum of constituent modes. The intensity field, which is related to the pressure field in a complex way, can be calculated for each mode. Boundary conditions relate averaged intensity and pressure. The mean-square pressure is expressed in terms of the modal sum. Lower order modes are responsible for the overall smooth spatial variation in the reverberant field; higher modes account for more rapid local variations near walls due to changes in wall properties. In the transient problem, the spatial eigenmodes decay at different rates, leading to a spatial redistribution of reverberant energy, and causing the decay curves to be a function of position in the room. Sample calculations and comparisons with other solution methods are presented.

9:50

5aAA4. An acoustical analysis of a room with a concave dome ceiling element. Sentagi S. Utami (N283 Eyring Sci. Ctr., Brigham Young Univ., Provo, UT 84602, sentagi@yahoo.com)

Concave surfaces are often considered detrimental in room acoustics, especially because of the impact they have on the distribution of sound energy. This paper explores certain acoustical characteristics and anomalies found in spaces below concave dome ceiling elements. The architectural design of the Darusshollah mosque in East Java, Indonesia is used as a case study with specific spatial and functional concerns. Investigations of the mosque have been conducted through both a 1:12 scale model and a computer model that utilizes ray tracing and image source methods. Analysis techniques are discussed. Results are presented and compared to provide useful insights into the acoustics of such distinctive environments.

10:05

5aAA5. Investigation on the flanking transmission of impact sound insulation of floor. Giovanni Semprini and Alessandro Cocchi (DIEMCA Dept. of Eng., Univ. of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy)

Impact sound pressure level of floors depends, as general rule, on direct sound radiated by the floor excited by a standard tapping machine and on flanking transmission of lateral walls. Depending on the kind of junction between the floor and walls of the receiving room, flanking paths can be more or less important. Requirements of laboratory test specimens are not well specified in EN ISO 140 standards, particularly for junctions of the test floor and lateral walls. In this paper measurements performed at DIENCA laboratory are presented in order to evaluate the influence of flanking transmission on impact sound pressure levels of a standard floor and on impact sound reduction level of standard floor with a resilient layer. Measurements are performed on a 14-cm-thick concrete floor in two different conditions: first connected on two sides of the receiving room and then on all four sides. Impact sound pressure levels and vibration levels are analyzed in order to evaluate the contribution of different transmission paths. As the measurements were carried on in the new facility for measurement of the contribution of the flanking transmission, this facility will be exhaustively presented in the paper.

9:25

5aAA6. Acoustics of early music spaces from the 11th to 18th century: Rediscovery of the acoustical excellence of medium-sized rooms and new perspectives for modern concert hall design. Alban Bassuet (Arup Acoust., 155 Ave. of the Americas, New York, NY 10013, alban.bassuet@arup.com)

The acoustical characteristics of 50 rooms that played a prominent role in the history of music between the 11th and 18th centuries were studied. The rooms include basilicas, oratorios, organ churches, and the great halls and courts of the European palaces. The research provides an understanding of the acoustical features that suit the early music repertoire, and how these rooms achieved an enhanced emotional engagement through their unique acoustical characteristics. This paper provides a summary of the acoustic measurements, which include binaural and B-format recordings in each of the rooms, and presents a unique new approach to understanding their subjective characteristics through detailed analysis and auralization of their 3-D impulse response. The study shows that the timing and direction of reflections in three dimensions is critically important to defining the subjective characteristic of a room. The results emphasize the importance of developing techniques to understand the 3-D impulse response and using auralization techniques for interpreting results and making subjective judgments. The enhanced musical experience that is achieved in these early rooms offers an invitation to rethink modern acoustics and to develop a new design approach that focuses more strongly on the subjective response and emotional engagement of the music.

10:35

5aAA7. Use of surrogate samples to study variation of absorption coefficients of fiberglass with altitude. Richard D. Godfrey (Owens Corning, Sci. and Technol., 2790 Columbus Rd., Granville, OH 43023)

ASTM C 423 identifies air temperature and relative humidity as significant parameters, but does not address air density effects. At constant temperature, air density decreases approximately 20% from sea level to 5000 ft altitude. In previous papers, normal and diffuse field analysis showed significant changes in predicted absorption coefficients with altitude. These predictions were validated experimentally for normal incidence by making measurements in a vacuum chamber. Reverberation chambers cannot withstand depressurization. They also exhibit significant interlaboratory measurement variability. Another method was soot. The Mechel design charts are normalized by two parameters. One is not dependent on air density. The other is the ratio of flow resistance and the impedance of air. If thickness is held constant, the effect of lowering air density can be studied by increasing the sample flow resistivity of the sample. This surrogate sample should emulate absorptive performance at high elevations in sea level laboratories. Impedance tube measurements using surrogate samples emulated the effects observed in the pressure chamber study. The next step is to use surrogate samples to investigate air density effects in diffuse fields using the ASTM C 423 test method in a single laboratory.

10:50

5aAA8. Acoustical phenomenon in ancient Totonac's monument. José Sánchez-Dehesa, Andreas Håkansson (Nanophotonic Technol. Ctr. and Dept. of Electron. Eng., Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), Francisco Cervera, Francisco Meseguer (Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), Betsabé Manzanares-Martínez, and Felipe Ramos-Mendieta (Univ. of Sonora, Hermosillo, Sonora 83190, Mexico)

The circle of gladiators is a monument built by Totonac Indians in the ceremonial site of Cempoala, which is located near Veracruz (Mexico). The city is believed to date to around 1200 A.D. The monument is a round structure with crenellated wall tops, and it has a diameter of 13.4 m. Though the deterioration of this monument is noticeable, it presents a singular acoustical phenomenon whose strength had to be probably extraordinary on the date of its construction. In brief, along any diameter in the circle, one can find two focal points such that if one person speaks on one focus, another person located on the other hears the sound reinforced. In other words, this circular place acoustically behaves as if it were elliptical. Here, we report the experimental characterization of the phenomenon and present a theoretical explanation. Also, the intentionality of the Totonacs is speculated since these people are associated with the Mayan culture, which is known by its realizations of environments with astonishing sonic properties. [Work supported by CEAL-UAM of Spain.]

FRIDAY MORNING, 28 MAY 2004 NEW YORK BALLROOM A, 8:00 TO 10:45 A.M.

Session 5aAOa

Acoustical Oceanography and Animal Bioacoustics: D. Van Holliday Special Session on Acoustical Measurements of Marine Organisms III

John K. Horne, Chair

School of Aquatic and Fisheries Science, University of Washington, Box 355020, Seattle, Washington 98195-5020

Contributed Papers

8:00

5aAOa1. Twenty-five years with Van Holliday in the development of high-frequency technology and analysis algorithms to measure zooplankton distributions. Richard E. Pieper (Southern California Marine Inst., 820 S. Seaside Ave., Terminal Island, CA 90731, pieper@usc.edu)

Initial studies using high-frequency acoustics at four individual frequencies (0.5–3.0 MHz) were begun in the 1970s to measure acoustical scattering from zooplankton. Acoustical measurements were made at sea by profiling vertically in the water column. Zooplankton were collected, identified and measured, and target strength measurements were made on individual zooplankton in the laboratory. Concurrently, various acoustical scattering models were analyzed to enable the calculation of the size-frequency distribution of zooplankton from the acoustical data. A multifrequency acoustic profiling system (MAPS) was then developed (21 different acoustical frequencies). This system was used to measure oceanic structure off of southern California, plumes off of central California, Gulf Stream features, and oceanic structure in the Irish Sea. Analyses of these data indicated that 21 frequencies were more than needed. Four to six frequencies were adequate for most studies, and the Tracor Acoustic Profiling System (TAPS) was developed. This system has been used and modified for a wide variety of studies. These studies range from large-scale patterns of zooplankton in the Arabian Sea to the measurement of thin layers in many different oceanic systems. The use of these systems now provides us with high-resolution measurements of zooplankton distributions in the sea.

8:15

5aAOa2. Acoustic backscatter models of fish: Gradual or punctuated evolution. John K. Horne (Univ. of Washington, School of Aquatic and Fishery Sci., Box 355020, Seattle, WA 98195, jhorne@u.washington.edu)

Sound-scattering characteristics of aquatic organisms are routinely investigated using theoretical and numerical models. Development of the inverse approach by van Holliday and colleagues in the 1970s catalyzed the development and validation of backscatter models for fish and zooplankton. As the understanding of biological scattering properties increased, so did the number and computational sophistication of backscatter models. The complexity of data used to represent modeled organisms has also evolved in parallel to model development. Simple geometric shapes representing body components or the whole organism have been replaced by anatomically accurate representations derived from imaging sensors such as computer-aided tomography (CAT) scans. In contrast, Medwin and Clay (1998) recommend that fish and zooplankton should be described by simple theories and models, without acoustically superfluous extensions. Since van Holliday's early work, how has data and computational complexity influenced accuracy and precision of model predictions? How has the understanding of aquatic organism scattering properties increased? Significant steps in the history of model development will be identified and changes in model results will be characterized and compared. [Work supported by ONR and the Alaska Fisheries Science Center.]

8:30

5aAOa3. A nonlinear model-based acoustic inversion to estimate the abundance and biomass distributions of marine organisms. Dezhang Chu and Peter Wiebe (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

Multi-frequency and/or broadband acoustic systems can be used to estimate the abundance and biomass distributions of marine organisms with inversion techniques [Holliday *et al.*, J. Cons. Int. Explor. Mer. **46**, 52–67 (1989)]. The linear inversion scheme is commonly used in fisheries and zooplankton acoustics and is based mostly on relatively simple scattering models. A more advanced nonlinear inversion method is presented in this talk. It uses the more sophisticated scattering models and the material properties measured *in situ* to estimate the abundance and biomass of marine organisms. A multi-frequency acoustic data set collected with the bio-optical multi-frequency acoustical and physical environmental recorder (BIOMAPER-II) during the austral fall Southern Ocean GLOBEC broad-scale cruise in 2002 was used in the inversion. At the depth where the Antarctic krill (*Euphausia superba*) were the dominate scatterers, the behavioral information of the animal such as mean angle of orientation and the standard deviation of the tilt angle were estimated with the nonlinear inversion. Furthermore, nonuniqueness and uncertainty inherently associated with the nonlinear inversion are analyzed. [Work supported by the NSF.]

8:45

5aAOa4. High-frequency acoustic volume scattering from zooplankton and moving oceanic microstructure. Andone C. Lavery, Peter H. Wiebe, Raymond W. Schmitt, Timothy K. Stanton, Tetjana Ross, Gareth Lawson, Nancy Copley (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Karen E. Fisher (Los Alamos Natl. Lab., Los Alamos, NM 87545), and Fabian Wolk (Rockland Oceanogr. Services, Inc., Victoria, BC V9A 4B6, Canada)

It is well accepted that high-frequency acoustic scattering techniques can be used to perform rapid, synoptic surveys of fish and zooplankton over relevant spatial and temporal scales. However, the use of these remote sensing techniques to probe small scale physical processes, such as oceanic microstructure, has not been fully accepted, or exploited. Yet there is a growing body of evidence in the form of both laboratory and field measurements suggesting that their use is feasible. Discrimination of scattering from microstructure versus zooplankton, which span similar spatial scales, is typically a limiting factor. Currently, acoustic discrimination of turbulence from zooplankton relies on either source of scattering being dominant as well as on the availability and accuracy of scattering models. A model for scattering from oceanic microstructure that includes fluctuations in the density and sound speed is presented. The effects on scattering from a layer of microstructure with a mean fluid velocity are discussed. Backscattering predictions are made based on data collected in Hudson Canyon with a tethered free-falling high-resolution vertical microstructure profiler, and compared to two-frequency acoustic data (120 and 420 kHz). The contribution to scattering from zooplankton is also estimated from nearby depth-resolved net tows together with zooplankton scattering models.

9:00

5aAOa5. Ocean acoustic backscattering: When you can ignore acoustic scatter from turbulence and when you can't. Tetjana Ross, Andone Lavery (Appl. Ocean Phys. and Eng., Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), Rolf Lueck (Univ. of Victoria, Victoria, BC V8W 2Y2, Canada), Peter Wiebe, and Gareth Lawson (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

While models predicting measurable levels of acoustic backscattering from oceanic turbulence have been around for decades, they have proven notoriously hard to confirm. This is, in part, because potential turbulent scattering layers often coincide with zooplankton layers. Therefore, as zooplankton are known to cause measurable acoustic scatter, the source of scatter is obscured. Furthermore, estimates of zooplankton abundance from acoustic scattering measurements have been shown to agree with independent measures under a number of circumstances. This gives circumstantial evidence to fuel the belief that scatter from turbulence is negligible. In addition, even if the turbulent scattering theory is correct, it predicts that over most of the ocean the turbulent intensities and/or stratifications are too weak to give turbulent scatter of a similar magnitude as is observed from zooplankton layers. Yet, despite all this, here data are presented that show circumstances when scatter from turbulence is significant. The areas of the ocean (mostly coastal) that one might have to be concerned about turbulent scatter are discussed, as well as the intriguing idea of using technology already developed for zooplankton observation to measure turbulence. This technology could facilitate the *in situ* study of biophysical interactions between zooplankton and turbulence.

9:15

5aAOa6. Effect of orientation of euphausiids and copepods on acoustic target strength: Implications for measurements from down-looking and side-looking acoustic systems. Malinda Sutor and Timothy J. Cowles (College of Oceanic and Atmospheric Sci., Oregon State Univ., 104 Ocean Admin. Bldg., Corvallis, OR 97331, msutor@coas.oregonstate.edu)

Multifrequency acoustics is a potentially useful tool for zooplankton ecologists to rapidly map distributional patterns and determine taxonomic and size composition of scatterers. This is typically done with down-looking or side-looking acoustic systems. The ability to estimate taxonomic and size composition acoustically depends upon accurate models of scattering from zooplankton. In this study, distorted wave Born approximation (DWBA) models of individual zooplankton were used to illustrate that the frequency dependence of expected scattering is different for down-looking and side-looking acoustic systems due to the differences in orientation of scatters within the sampled volume. The results show that peaks and nulls occur at different frequencies for each system with maximum target strength differences of 30 dB. The models were used to predict volume scattering (S_V) based on zooplankton collected from MOCNESS tows. Predicted S_V was compared to S_V measured by both a down-looking HTI acoustic system and a side-looking TAPS. The results showed that changes in orientation can have large effects on predicted total S_V, particularly at higher frequencies, and demonstrate that choice of orientation parameters is important and different parameters should be used when comparing predicted S_V with measured S_V from down-looking or side-looking acoustic systems.

9:30

5aAOa7. Comparing high-frequency scattering by a fish swimbladder and a gas-filled ellipsoid. Kenneth G. Foote (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543) and David T. I. Francis (Univ. of Birmingham, Birmingham B15 2TT, UK)

High-frequency backscattering spectra have been computed for two air-filled sacs at 51-atm ambient pressure by means of the boundary-element method. One applies to the swimbladder of a 39-cm-long specimen of pollack (*Pollachius pollachius*), with known morphometry expressed in a 3181-node mesh. The other applies to a prolate ellipsoid, whose major axis is that of the mapped swimbladder, 14.1 cm, and whose minor axis, 1.24 cm, has been determined so that the volume of the ellipsoid is essentially identical to that of the mapped swimbladder. The dorsal-aspect backscattering cross section of each sac has been computed at each frequency over a normal distribution of tilt angles, measured relative to the longitudinal or major axis, with mean −4.4 deg and standard deviation 16 deg. Computations have been performed over the frequency range 20–40 kHz in increments of 25 Hz. The spectra are characterized and compared. [Work supported by ONR.]

9:45

5aAOa8. Some anomalous time domain back scattering phenomena from low contrast fluid spheres and cylinders. C. Feuillade (Naval Res. Lab., Stennis Space Ctr., MS 39529-5004), D. Chu (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543), and C. S. Clay (Geophys. and Polar Res. Ctr., Univ. of Wisconsin—Madison, Madison, WI 53706)

Anderson's scattering theory [J. Acoust. Soc. Am. **22**, 426–431 (1950)], and the corresponding formalism for an infinite cylinder, have been used to calculate the scattered impulse responses for low contrast fluid spheres and cylinders, where the density and sound speed of the object are fractionally higher than the surrounding medium. The returns from the front and rear faces of the object are individually identifiable. In the case of a sphere, both of these features are positive spikes. In the cylindrical case, the second feature is a negative first derivative of a spike. In both instances, the waveform of the second feature is different from that expected from simple physical reasoning, i.e., the raypaths and reflection coefficients. The predictions of the calculations are investigated by using them to analyze experimental scattering data from decapod shrimp near broadside incidence. By convolving the calculated impulse response with the incident acoustic signal, comparisons with the scattering data are made. Both the sphere and cylindrical calculations appear to fit the data well. The significance of these results, and possible explanations for the anomalous scattering from the rear face, are discussed. [Work supported by ONR, University of Wisconsin Weeks Fund.]

10:00

5aAOa9. Modeling of surficial sediment alteration by biology. Dajun Tang (Appl. Phys. Lab., Univ. of Washington, Seattle, WA 98105)

This paper describes fine-scale measurements and modeling of surficial sediment roughness at a site in the East China Sea using a conductivity probe. The spatial resolution of the measurements is designed to obtain environmental data suitable for modeling acoustic backscatter in the mid-frequency (3–4 kHz) range. The power spectrum of the bottom roughness is estimated and it is found that bottom roughness is dominated by small features caused by bottom-dwelling organisms. This is confirmed by video images of the same spot of seafloor. A model is developed to simulate the random distributions of these bottom features. The model employs a superposition of discrete features, which result in a power spectrum that is consistent with the measured power spectrum. Potentially this kind of model can provide a remote sensing means to estimate bottom biological populations through measuring sound backscatter from the bottom. [Work supported by ONR Ocean Acoustics Code.]

10:15

5aAOa10. The Bergen multifrequency analyzer (BMA): A new toolbox for acoustic categorization and species identification. Egil Ona, Rolf Korneliussen, Hans Petter Knudsen (Inst. of Marine Res., P.O. Box 1870, 5817 Bergen, Norway, egil.ona@imr.no), Kjell Rang, Inge Eliassen, Yngve Heggelund, and Daniel Patel (Christian Michelsen Res. AS, 5892 Bergen, Norway)

Multifrequency split-beam echo sounders with nearly identical and overlapping acoustic beams have been regularly used in acoustic surveys for fish stock abundance estimation. Calibrated raw data from up to six simultaneously working echo sounders at 18, 38, 70, 120, 200, and 364 kHz were applied for developing a new processing tool for real-time acoustic target categorization and acoustic species identification. The system now handles raw data from the Simrad EK500 and EK60 split-beam echo sounders, and performs a stepwise, modular sequence of analysis, like bottom detection, noise quantification and removal, target categorization, and school detection in near-real time. Direct generation of new, synthetic echograms, based upon the measured frequency response of the targets, is also one of the most useful features of the system. This information may significantly increase the accuracy of acoustic survey estimates of fish and zooplankton. New routines for noise removal, target categorization, and school detection will be presented, as well as new methods for training and building the artificial experience of the analyzer.

10:30

5aAOa11. An echo analysis technique for estimating the fish population. C. P. Anil Kumar, Sajith N. Pai, N. Soniraj, M. H. Supriya, James Kurian, C. Madhavan, and P. R. Saseendran Pillai (Dept. of Electron., Cochin Univ. of Sci. and Technolgy, Cochin-22, Kerala, India)

The development of an algorithm for the estimation of biomass by acoustic remote sensing is presented in this paper. The distinctive features of the algorithm include the implementation of time-varied gain function, proper accounting of beam factor effects, implementation of backscattering levels for selected species and processing for echo-count as well as echo-integration. In order to optimize the backscattering levels, numerical analysis for geometrical backscattering of selected marine species were carried out and validated with *in situ* measurements. The biomass information of the concerned marine specie is made available by subjecting the backscattered raw data to a series of processes. The performance validation of the algorithm under *in situ* conditions yielded encouraging population estimation results and is being fine-tuned with the field data. This algorithm will provide an efficient technique to parametrically compute the target strength, leading to the estimation of the stock of commercially important selected marine species under varied environmental conditions in different regions.

Session 5aAOb

Acoustical Oceanography: Acoustical Oceanography Prize Lecture

Peter F. Worcester, Chair
Scripps Institution of Oceanography, University of California—San Diego, 9500 Gilman Drive, La Jolla, California 92093-0225

Chair's Introduction—10:55

Invited Paper

11:00

5aAOb1. Ocean acoustic inversion for seabed geoacoustic properties. Stan E. Dosso (School of Earth and Ocean Sci., Univ. of Victoria, Victoria, BC V8W 3P6, Canada, sdosso@uvic.ca)

Estimating geoacoustic properties of the seabed from ocean acoustic data provides a convenient *in situ* alternative to direct sampling (e.g., coring), with parameter sensitivities relevant to sonar applications. However, this requires solving a strongly nonlinear inverse problem, which is inherently nonunique. Hence, quantifying uncertainties for the recovered geoacoustic parameters is an important, but challenging, problem. This talk will describe a nonlinear Bayesian approach to geoacoustic inversion based on estimating properties of the posterior probability density (PPD), which combines information from observed data with prior information. An efficient Markov-chain method (Gibbs sampling) is applied to extract properties of the PPD, including optimal parameter estimates, marginal probability distributions, variances/covariances, and interparameter correlations. The inversion formulation is general, and will be illustrated with examples for a variety of approaches to geoacoustic inversion, including the inversion of acoustic field data, seabed reflectivity measurements, acoustic reverberation data, and ambient noise measurements.

Session 5aBB

Biomedical Ultrasound/Bioresponse to Vibration: Tissue Characterization and Gene Transfection

Subha Maruvada, Chair
FDA Center for Devices and Radiological Health, Rockville, Maryland 20852

Contributed Papers

8:00

5aBB1. Acquisition of transcranial ultrasonic signals from adult subject's applications to the detection of brain injury. Joel Mobley, Tuan Vo-Dinh (Oak Ridge Natl. Lab., P.O. Box 2008, Oak Ridge, TN 37831-6101, vodinht@ornl.gov), Brian J. Daley (Univ. of Tennessee Medical Ctr., Knoxville, TN 37920), and Martin C. Holland (San Francisco General Hospital, San Francisco, CA 94143)

In the clinical setting, several technologies (e.g., x-ray CT, MRI, color-duplex ultrasonography) can provide doctors with noninvasive windows through the cranium into the brain. In the field, however, there is a need for a portable device that can noninvasively detect brain injuries at the scene of a trauma, providing first responders with urgent diagnostic information in the critical first hour postincident. Such an instrument should be simple to operate and must supply results that are straightforward to interpret. This work is part of an effort to devise ultrasonic methods that can serve as the operating principle for such a device. In this talk, human studies are described in which backscattered ultrasound data were obtained transcranially from adult volunteers. The ultrasound signals, acquired with single transducers in pulse–echo mode, are interpreted anatomically with the aid of a multilayer model of the propagation path. Through the model, the complex structures of the echo patterns are revealed and the important roles of the scalp and skull layers are made evident. The ultrasound data are further analyzed to assess the suitability of these signals for implementing a tissue characterization approach to the injury detection problem.

8:15

5aBB2. Importance of *ka*-range on the simultaneous estimation of scatterer size and total attenuation from ultrasound backscattered waveforms. Timothy A. Bigelow and William D. O'Brien, Jr. (Dept. of Elec. and Computer Eng., Bioacoustics Res. Lab., Univ. of Illinois, Urbana, IL 61801, bigelow@uiuc.edu)

Considerable effort has been directed to diagnosing the malignancy of solid tumors noninvasively by determining a scatterer size from a statistical analysis of the ultrasound-backscattered waveforms. *In vivo* tumor assessments have had limited success due to frequency-dependent attenuation along the propagation path masking the frequency dependence of the scatterer size. In this study, both attenuation and size were solved simultaneously by a two-parameter minimization of the mean squared error between a reference spectrum, modified by the attenuation and scatterer size, and the backscattered waveforms. The performance of the approach was assessed by simulations of a homogeneous region containing scatterers with a Gaussian impedance distribution. The simulations varied the effective radius of the scatterers (5 to 150 mm), the attenuation of the

region (0 to 1 dB/cm/MHz), and the bandwidth of the source. In all cases, comparable accuracy and precision of the scatterer size were obtained whenever the range of *ka* values (wavenumber times effective radius) had the same width (largest *ka* minus smallest *ka*). The precision and accuracy improved with increasing width. A width of 1 gave an accuracy/precision of $\sim 15\% \pm 35\%$ whereas a width of 1.5 gave an accuracy/precision of $\sim 5\% \pm 15\%$. [Work supported by a Beckman Institute Fellowship.]

8:30

5aBB3. Characterization and differentiation of three solid tumors using quantitative ultrasound. Michael L. Oelze, William D. O'Brien, Jr. (Dept. of Elec. and Computer Eng., Univ. of Illinois, 405 N. Mathews, Urbana, IL 61801), and James F. Zachary (Univ. of Illinois, Urbana, IL 61801)

Three kinds of solid tumors were acquired and scanned *in vivo* ultrasonically. The first tumor series (fibroadenoma) was acquired from tumors that had spontaneously developed in rats. The second tumor series was acquired by culturing a carcinoma cell line (4T1-MMT) in culture media and injecting the cells into Balb/c mice. The third tumor was acquired by transplanting a soft-tissue sarcoma cell line (EHS) into C57BL mice. The tumors were allowed to grow to 1 cm in size and then scanned ultrasonically. The scatterer properties of average scatterer diameter and acoustic concentration were estimated using a Gaussian form factor from the backscattered ultrasound measured from the tumors. Parametric images of the tumors were constructed utilizing estimated scatterer properties for regions of interest inside the tumors. The parametric images showed distinct differences between the various tumor types. Quantitatively, the tumors could be distinguished through feature analysis plots of average scatterer size versus acoustic concentration. Comparison with photomicrographs of the tumors showed structures similar in size to the ultrasound estimates. [Work supported by NIH Grant F32 CA96419 to MLO and by the University of Illinois Research Board.]

8:45

5aBB4. Evaluation of the DORT method for the detection of microcalcifications in the breast. Jean-Luc Robert, Claude Cohen-Bacrie (Philips Res. USA, 345 Scarborough Rd., Briarcliff Manor, NY 10510-2099), Claire Prada, and Mathias Fink (Universite Denis Diderot, 75231 Paris Cedex 05, France)

The DORT method (French acronym for diagonalization of the time reversal operator) is derived from the theory of iterative time reversal mirroring. It consists of a singular value decomposition of the time reversal operator obtained through single element transmissions and receptions. The number of eigenvalues relates to the number of bright point scatterers in the medium, and each eigenvector is the transmit signals that focuses on each bright scatterer. However, the signal-to-noise ratio (SNR) resulting from a single element transmit is low, which negatively impacts the sensitivity of the DORT method. This work consists of an adaptation of the DORT method to an imaging mode. Focused transmissions in the medium are used and a windowing preprocessing operation on the received signals significantly increases the sensitivity. The more robust behavior of this modified DORT method is tested on Field II simulated data and then on a phantom made of strings of different material embedded in speckle. Data on freshly excised breast surgical samples containing microcalcifications were processed and every microcalcifications was detected. Finally, results on *in vivo* acquisitions prove that the technique can accurately detect and locate microcalcifications in nonhomogeneous media with aberration.

9:00

5aBB5. Scatterer size and concentration estimation technique based on a 3D acoustic impedance map from histologic sections. Jonathan Mamou, Michael L. Oelze, William D. O'Brien, Jr. (Dept. of Elec. and Computer Eng., Univ. of Illinois, 1406 W. Green St., Urbana, IL 61801), and James F. Zachary (Univ. of Illinois, Urbana, IL 61801)

Accurate estimates of scatterer parameters (size and acoustic concentration) are beneficial adjuncts to characterize disease from ultrasonic backscatterer measurements. An estimation technique was developed to obtain parameter estimates from the Fourier transform of the spatial autocorrelation function (SAF). A 3D impedance map (3DZM) is used to obtain the SAF of tissue. 3DZMs are obtained by aligning digitized light microscope images from histologic preparations of tissue. Estimates were obtained for simulated 3DZMs containing spherical scatterers randomly located: relative errors were less than 3%. Estimates were also obtained from a rat fibroadenoma and a 4T1 mouse mammary tumor (MMT). Tissues were fixed (10% neutral-buffered formalin), embedded in paraffin, serially sectioned and stained with H&E. 3DZM results were compared to estimates obtained independently against ultrasonic backscatter measurements. For the fibroadenoma and MMT, average scatterer diameters were 91 and 31.5 μm, respectively. Ultrasonic measurements yielded average scatterer diameters of 105 and 30 μm, respectively. The 3DZM estimation scheme showed results similar to those obtained by the independent ultrasonic measurements. The 3D impedance maps show promise as a powerful tool to characterize ultrasonic scattering sites of tissue. [Work supported by the University of Illinois Research Board.]

9:15

5aBB6. Ultrasonic tissue characterization for the classification of prostate tissue. Ulrich Scheipers, Helmut Ermert (Lehrstuhl fuer Hochfrequenztechnik, Ruhr-Universitaet Bochum [RUB], Germany, ulrich.scheipers@rub.de), Katharina Koenig, Hans-Joerg Sommerfeld, Miguel Garcia-Schuermann, Theodor Senge (Urologische Universitaetsklinik der RUB, Marienhospital Herne, Germany), and Stathis Philippou (Institut für Pathologie, Augusta-Krankenanstalt Bochum, Germany)

Radio-frequency ultrasound echo data of the prostate are captured during routine examinations with standard ultrasound equipment. The data are directly transmitted to a PC and subdivided into numerous regions of interest. Several parameters describing the histological characteristics of the underlying tissue are calculated from the frequency spectrum and from the demodulated signal of the underlying echo data. Parameters are fed into two adaptive network-based fuzzy inference systems working in parallel. One system is used to classify hypo- and hyperechoic tumors, the other system is used to detect isoechoic tumors within the normal prostate tissue. Subsequent morphological analysis combines clusters to mark areas of similar tissue characteristics. Classification results are presented as two-dimensional malignancy maps and as volumetric reconstructions of the whole organ. Radio-frequency ultrasonic echo data of 100 patients have been recorded. Tissue samples following radical prostatectomies are used as the gold standard. The area under the ROC curve is $A = 0.86 \pm 0.01$ for hypo- and hyperechoic tumors and $A = 0.84 \pm 0.02$ for isoechoic tumors using leave-one-out cross validation over patient datasets.

9:30

5aBB7. Sonoporation of cell using therapeutic ultrasound for drug and gene delivery. Hua Pan, Fred Sieling, Yun Zhou, Hesheng Wang, Jianmin Cui, and Cheri Deng (Dept. of Biomed. Eng., Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106-7207)

Recent studies of ultrasound methods for targeted drug delivery and nonviral gene transfection revealed new, advantageous possibilities. These studies utilized ultrasound contrast agents, commonly stabilized microbubbles, to facilitate delivery and suggested that ultrasound delivery resulted from cell sonoporation, the formation of temporary pores in the cell membrane induced by ultrasound. In this study, voltage clamp techniques were used to obtain real-time measurements of sonoporation of single Xenopus oocytes in the presence of Optison. Ultrasound increased the transmembrane current as a direct result of decreased membrane resistance due to pore formation. We observed a distinct delay of sonoporation following ultrasound activation and characteristic stepwise increases of transmembrane current throughout ultrasound duration. We discovered that the resealing of the cell membrane following ultrasound exposure required calcium entering the cell through ultrasound-induced pores.

9:45

5aBB8. A comparison study in cell transfection: Which one is better, sonoporation versus electroporation? Junru Wu, Jason Pepe (Dept of Phys., Univ. of Vermont, Burlington, VT 05405), and Mercedes Rincon (Univ. of Vermont, Burlington, VT 05405)

An experimental study has been performed for cell suspensions to compare efficiency of cell transfection, which is the process of introducing recombinant DNA into eukaryotic cells (eukaryotic cells have chromosomes with nucleosomal structure) and subsequently integrating that DNA into the recipient cell's chromosomal DNA. It was demonstrated that electroporation was superior to sonoporation in terms of viability (65.8 2.3% vs. 50.8 4.15%) and transfection efficiency (15.83 3.5% vs. 7.53 0.4%) for Jurkat lymphocytes (nonprimary cells), and sonoporation was better in terms of viability (64.8 1.51% vs. 53.7 1.53 %) and transfection efficiency (2.73 0.21% vs. 0.43 0.06%) for human peripheral blood mononuclear cells (primary cells). The electroporation was performed using a Gene Pulser II Apparatus with voltage of 250 V, and the sonoporation was achieved using 2-MHz pulsed ultrasound exposure (ISPPA = 80 W/cm^2) assisted with encapsulated bubbles (Optison).

10:00

5aBB9. HIFU-induced gene activation *in vitro*. Yunbo Liu, Pei Zhong (Dept. of Mech. Eng., Duke Univ., Durham, NC 27708), Takashi Kon, and Chuanyuan Li (Duke Univ., Durham, NC 27708)

This work investigated the inducible gene activation in cancer cells that were sublethally injured during HIFU treatment. HeLa cells were transfected by an adenovirus vector that encodes GFP under the control of hsp70B promoter, leading to about 65% transfection efficiency. A volume of 10 μL transfected HeLa cells in suspension (5×10^7 cells/ml) were placed at the bottom of a PCR tube so that the cell suspension could be heated to a peak temperature of 50 °C, 60 °C, and 70 °C for 120, 10, and 1 s, respectively, by a focused 1.1-MHz HIFU transducer operated at a peak negative pressure of -2.7 MPa at different duty cycles. One day after HIFU treatment, cell viability was determined to be 63%, 35%, and 18%, respectively, based on Trypan Blue exclusion test. Importantly, in all test groups, inducible GFP expression was detected in about 40%–50% of the surviving cells with GFP intensity increased by 25-fold based on flow cytometry analysis. These results demonstrate that even under the short exposure duration of HIFU treatment, inducible gene expression could be produced in sublethally injured cell population *in vitro*. Further studies are underway to explore the optimal HIFU condition for gene activation *in vivo*.

10:15

5aBB10. Why does humming clean your maxillar sinuses of NO gas? Johan Sundberg, Svante Granqvist (Speech Music Hearing, KTH, SE-100 44 Stockholm, Sweden, pjohan@speech.kth.se), Eddie Weitzberg, and Jon Lundberg (Karolinska Inst., Karolinska SE-17177, Sweden)

Recent measurements have shown that the nitric oxide (NO) produced in the maxillar sinuses can be evacuated by producing nasal murmur; the NO content of the nasal airflow increases under these conditions (Maniscalco *et al.*, 2003). Experiments were carried out to test the hypothesis that this effect is caused by resonance in the nasal tract and sinus cavities. A model was constructed where an airflow from a pressure tank was modulated at different frequencies. This airstream was passed through a tube with a radial hole constituting the neck of a Helmholtz resonator with a gas containing NO. The NO content of the air streaming out of the tube was measured. This NO content varied when the location of the resonator, its air volume, or the modulation frequency of the airflow was changed. The relevance of these three factors was also tested by means of a computer model of the system. The significance of the modulation frequency was also analyzed in human subjects by injecting a modulated airflow through one nostril and measuring NO content of the airflow exiting the other nostril, the results showing a dependence on the modulation frequency that differed considerably between subjects [Maniscalco *et al.*, ''Assessment of nasal and sinus nitric oxide output using single-breath humming exhalations,'' Eur. Respir. J. **22**, 323–329 (2003)].

FRIDAY MORNING, 28 MAY 2004 LIBERTY 5, 8:25 TO 11:15 A.M.

Session 5aEA

Engineering Acoustics: Computational Acoustics, Ultrasonics and Applications

Elizabeth A. McLaughlin, Cochair
Naval Undersea Warfare Center, 1176 Howell Street, Newport, Rhode Island 02841

Juan Arvelo, Cochair
Applied Physics Laboratory, Johns Hopkins University, 11100 Johns Hopkins Road, Laurel, Maryland 20723-6099

Chair's Introduction—8:25

Contributed Papers

8:30

5aEA1. A Galerkin discontinuous method for CAA: The reasons of a choice. Philippe Delorme and Christophe Peyret (ONERA, BP 72, 92322 Chatillon, France)

As for many others disciplines, the numerical simulation is becoming a powerful tool for the study of the propagation of small disturbances in a heterogeneous flow (aero-acoustics). Some main applications are, on the one hand, aeronautics for the noise of the aircraft and, on the other hand, the propagation of the sound in the atmosphere. In general the development of a computer code goes by four phases: physical modeling, mathematic analysis, strategy of discretization, programming and validation. At each of these stages some specific choices are carried out. For physical modeling, it is necessary to choose a representative model. Then, it should be checked that the problem is well posed in existence and unicity. For the strategy of discretization the choice resides between the different methods and is guided by the geometries to study. Finally, the choice to use or not the possibility of strongly paralleling influences programming of the code. In this paper the reasons are described which led us to use a method of discontinuous the Galerkin type for the CAA. This method is then presented as well as architecture of the code. Finally, all this is illustrated for examples of simulations.

8:45

5aEA2. Approaches to simulate electromechanical coupling in voided piezoelectric materials. Juan Arvelo, Jr. (Johns Hopkins Univ., Appl. Phys. Lab., 11100 Johns Hopkins Rd., Laurel, MD 20723-6099) and Ilene Busch-Vishniac (Johns Hopkins Univ., Baltimore, MD 21218-2681)

Recent work has demonstrated piezoelectric behavior in voided polymers. It is thought that the piezoelectricity results from the creation of space charges bordering the voids when the material is exposed to an electric field in excess of the breakdown strength. The resulting voided piezoelectric material is more flexible than conventional piezoelectric ceramics, and may be far less expensive to manufacture. Accurate models of piezoelectricity in voided materials are needed to explore the limits of performance and their use under a wide range of conditions. We are using finite element analysis to develop a model of the mechanical, electrical, and piezoelectric behavior in voided materials. Results for some of the materials of current interest, such as low-density polypropylene (LDPP) and porous polytetrafluoroethylene (PTFE) will be presented. [Work performed under a JHU/APL sabbatical to the JHU/WSE.]

9:00

5aEA3. Simulation of linear aeracoustic propagation in lined ducts with discontinuous Galerkin method. Christophe Peyret and Philippe Delorme (ONERA BP, 72 Chatillon, France)

The simulation of the acoustic propagation inside a lined duct with nonuniform flow still presents problems when geometry is complex. To handle computations on a complex geometry, without consequential effort, unstructured meshes are required. Assuming irrotational flow and acoustic perturbation, a well-posed finite element method based on the potential equation is established. But, the effect of the thin boundary layer is then neglected, which is not relevant to the acoustical processes occurring near the lining. Recent works have focused on the tremendous interest of the Galerkin discontinuous method (GDM) to solve Euler's linearized equations. The GDM can handle computations on unstructured meshes and introduces low numerical dissipation. Very recent mathematical works have established, for the GDM, a well-posed boundary condition to simulate the lining effect. Results computed with the GDM are presented for a uniform cross section lined duct with a shear flow and are found to be in good agreement with the modal analysis results, thereby validating the boundary condition. To illustrate the flexibility of the method other applications dealing with instabilities, air-wing diffraction, and atmospheric propagation are also presented.

9:15

5aEA4. Engineering acoustic lenses with help from evolution. Andreas Håkansson, José Sánchez-Dehesa, and Lorenzo Sánchis (Nanophotonic Technol. Ctr. and Dept. of Electron. Eng., Polytechnic Univ. of Valencia, Spain)

Optimization engineering through evolutionary algorithms have proven to be very efficient, especially in hard problems containing a large set of optimization parameters. Like evolution this family of algorithms is able to tackle enormous complex problems with fairly simple means. Here, a simple genetic algorithm [J. H. Holland, *Adaptation in Natural and Artificial Systems* (Univ. of Michigan, Ann Arbor, 1975)] is used in conjunction with the multiple scattering theory [L. Sánchis *et al.*, Phys. Rev. B **67**, 035422 (2003)] to fabricate a new generation of acoustic devices based on a discrete number of cylindrical scatterers. In particular, acoustic lenses [F. Cervera *et al.*, Phys. Rev. Lett. **88**, 023902 (2002)] with flat surfaces have been designed to focus the sound in a fixed focal point for one or multiple frequencies. Each scatterer is carefully placed using the optimization method within the preset boundary conditions, to maximize the pressure contribution in the chosen focal spot. With this method acoustic lenses with very low f-numbers of the order 0.3 and with amplifications over 12 dB have been estimated using a reduced number of scatterers ($\sim$60). Preliminary results obtained from the experimental realization of the designed devices confirm our predictions.

9:30

5aEA5. An application of boundary element method calculations to hearing aid systems: The influence of the human head. Karsten B. Rasmussen (Oticon A/S, Strandvejen 58, DK-2900 Hellerup, Denmark) and Peter Juhl (Univ. of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark)

Boundary element method (BEM) calculations are used for the purpose of predicting the acoustic influence of the human head in two cases. In the first case the sound source is the mouth and in the second case the sound is plane waves arriving from different directions in the horizontal plane. In both cases the sound field is studied in relation to two positions above the right ear being representative of hearing aid microphone positions. Both cases are relevant for hearing aid development. The calculations are based upon a direct BEM implementation in Matlab. The meshing is based on the original geometrical data files describing the B&K Head and Torso Simulator 4128 combined with a 3D scan of the pinna.

9:45

5aEA6. Dissipative silencers with an extended inlet/outlet and baffles. Ahmet Selamet, Iljae Lee, Mubing Xu (The Ohio State Univ., 930 Kinnear Rd., Columbus, OH 43210, selamet.1@osu.edu), and Norman Huff (Owens Corning Automotive, Novi, MI 48377)

The acoustic characteristics of a single-pass perforated dissipative silencer were investigated experimentally and numerically by Selamet *et al.* [J. Acoust. Soc. Am. **109**, 2364 (2001)]. The current study extends this work by considering variations in the internal structure of the dissipative silencer. In addition to the boundary element method (BEM) introduced earlier, a multi-dimensional analytical approach is now developed to investigate the wave modes and transmission loss. Both methods are then employed to study the effect of an extended inlet and outlet on the acoustic behavior of the silencer. BEM is further used to explore the effect of baffles and air space inside the dissipative chamber. The location and number of baffles inside the dissipative chamber are shown to have a significant influence on the transmission loss.

10:00

5aEA7. Acoustic waves in discrete media, similarities at meso and nano scales. Hasson Tavossi (Dept. of Physical and Environ. Sci., Mesa State College, School of Math. and Physical Sci., 1100 North Ave., Grand Junction, CO 81501)

In this paper the similarities between the acoustic behavior of a discrete medium of random arrangement of solid spheres, at mesoscopic scales, and thermal vibrations of lattice ions in crystalline solid, at nano scales, are investigated. Results for the ultrasonic waves in random media show that such effects as cutoff frequency, wave dispersion, energy distribution in vibration modes, wave attenuation by scattering and absorption, observed in discrete media, have close resemblance to the similar phenomena observed at atomic scales, such as phonons or lattice thermal vibration of atoms in the crystalline solids. For example, the cutoff frequency in lattice vibration is related to the interatomic spacing, similarly, the cutoff frequency at mesoscopic scales depends on grain separation or grain size. These similarities are also observed in wave scattering and attenuation and their dependence on wave number, kR ($k=2\pi/\lambda$ and R is particle size). In this paper experimental measurements with data analysis on cutoff frequency, wave attenuation, and wave dispersion related to the above mentioned similarities will be presented, and the extension of these findings to the behavior of acoustic waves in other discrete media will be discussed.

10:15

5aEA8. Ultrasonic imaging in noisy environment. John Lewis, Matt Kaiser, James Irwin, Jr., and Jose Sanchez (Dept. of Elec. & Computer Eng., Bradley Univ., 1501 W. Bradley Ave., Peoria, IL 61625, jhirwin@bradley.edu)

Often times fossils are found buried within rocks of similar composition to the fossil. Ultrasonic imaging is becoming a popular form of nondestructive testing. The presentation covers techniques for imaging a fossil through nondestructive testing in three dimensions without risking damage to the fossil by removing extraneous material. The lecture begins with an illustration of immersion testing that uses longitudinal transducers to construct images of objects buried within a noisy environment of similar composition. Preliminary results give the location and shape of an object buried under a homogeneous layer. The first step to attaining this goal is to merely locate an object buried in a noisy debris field. A single sensor is all that is required to locate an object; however, two sensors allow the object's position to be triangulated. Techniques for using single and multiple sensors are investigated to improve the search and imaging algorithms. The lecture continues with an in-depth discussion of the MATLAB signal-processing algorithms, which are required to generate a visual representation of the object. With some modification this technique also promises to be useful for security screening and landmine detection.

10:30

5aEA9. Time reversal interactive objects. Ros Ki Ing (Sensitive Object, Res. and Development, 10 rue Vauquelin, 75005 Paris, France), Nicolas Quieffin, Stefan Catheline, and Mathias Fink (Paris 7 Univ., 75231 Cedex, Paris)

Time reversal has shown to be a fruitful concept in nondestructive testing in underwater acoustic or in ultrasonic imaging. In this paper this technique is adapted in the audible range to transform every day objects into tactile sensitive interfaces. A quick historical background is presented in the ultrasonic field and specially in chaotic cavity. In all time reversal experiments, it is demonstrated that a wave field spatially and temporally recorded is able to back propagate to its source. In other words, the field contains all the information on the location of the source. In the interactive experiments, it is shown that touching an object like a window, a table or a world globe generates an acoustic field easily detectable with one or two acoustic sensors. Using the concept of time reversal, the source location is deduced in real time. Then, touching objects at specific locations (virtual switches) is used to activate devices. Such devices are for example lights, stereo volume, or computer software. From a technical point of view, all these interactive experiments just use some computation easily performed with a standard personnel computer.

10:45

5aEA10. High-sensitivity photoacoustic leak testing. Eric Huang, David R. Dowling (Dept. of Mech. Eng., Univ. of Michigan, Ann Arbor, MI 48109-2133), Timothy Whelan (Honeywell FM&T LCC, Kansas City, MO 64141-6159), and John L. Spiesberger (Univ. of Pennsylvania, Philadelphia, PA 19104-6316)

The photoacoustic effect may be exploited for detection and localization of gas leaks on the surface of otherwise sealed components. The technique involves filling the test component with a photoactive tracer gas, and irradiating the component to produce photoacoustic sound from any leak site where a tracer gas cloud forms. This presentation describes demonstration experiments utilizing 10.6-μm radiation from a nominally 145-W carbon-dioxide laser with sulfur hexafluoride as a tracer gas. Here, photoacoustic sounds from six NIST-traceable calibrated leak sources with leak rates between 1 cc in 4.6 h, and 1 cc in 6.3 years were recorded with 12 microphones in a bandwidth from 3 to 80 kHz. Bartlett matched-field processing of the microphone array measurements both detect and localize these leaks when the leak and the array are separated by 152 mm. These experiments suggest that the sensitivity of photoacoustic leak testing may reach or even exceed the capabilities of the most sensitive commercial leak test systems using helium mass-spectrometers. Comparison of the measured results and an analytical scaling law suggests that tracer cloud geometry influences the photoacoustic signal amplitude. [Work supported by the U.S. Dept. of Energy.]

11:00

5aEA11. Evaluation of electromagnetic field interference on sound-pressure level measurements. Luigi Maxmilian Caligiuri and Adolfo Sabato (Dept. of Mech., Faculty of Eng., Univ. of Calabria—Via P. Bucci 87030 Arcavacata di Rende [CS], Italy, maxmc@tiscali.it)

The recently published IEC 61672-1 defines new standards for sound-level meters and calibrators, giving performance specifications, environmental and electromagnetic criteria, in order to overcome some important faults of the previous IEC 651-1979 and IEC 804-1985. Both of these standards have been in fact amended to include, in particular, specifications for immunity of electrostatic and electromagnetic fields. The interference of such nonacoustics factors can represent, if nonadequately recognized and quantified, an important uncertainty source in sound-pressure level measurements, able to cause errors on measurement result, variable between fractions of dB and some dB. In the present paper we will show the application of a measurement methodology able to quantify, within the framework of IEC 61672-1, the influence of electromagnetic fields on environmental noise measurements, carried out by a class I (according to IEC 651 and 804) integrating sound-level meter. In this way, we have analyzed the effect of different typologies of interfering field sources (characterized by different spectrum and emission features) on some very frequent noise-measurement configuration.

Session 5aMU

Musical Acoustics and Psychological and Physiological Acoustics: Neurophysiology of Playing a Musical Instrument

Ingo R. Titze, Chair
Department of Speech Pathology and Audiology, University of Iowa, 330 WJSHC, Iowa City, Iowa 52242-1012

Chair's Introduction—9:00

Invited Papers

9:05

5aMU1. Movement amplitude and tempo change in piano performance. Caroline Palmer (Dept. of Psychol., McGill Univ., 1205 Dr. Penfield Ave., Montreal, QC H3A 1B1, Canada, caroline.palmer@mcgill.ca) and Simone Dalla Bella (Kazimierz Wielki Univ., Bydgoszcz, 85–867 Poland)

Music performance places stringent temporal and cognitive demands on individuals that should yield large speed/accuracy tradeoffs. Skilled piano performance, however, shows consistently high accuracy across a wide variety of rates. Movement amplitude may affect the speed/accuracy tradeoff, so that high accuracy can be obtained even at very fast tempi. The contribution of movement amplitude changes in rate (tempo) is investigated with motion capture. Cameras recorded pianists with passive markers on hands and fingers, who performed on an electronic (MIDI) keyboard. Pianists performed short melodies at faster and faster tempi until they made errors (altering the speed/accuracy function). Variability of finger movements in the three motion planes indicated most change in the plane perpendicular to the keyboard across tempi. Surprisingly, peak amplitudes of motion before striking the keys increased as tempo increased. Increased movement amplitudes at faster rates may reduce or compensate for speed/accuracy tradeoffs. [Work supported by Canada Research Chairs program, HIMH R01 45764.]

9:25

5aMU2. Physiology, anatomy, and plasticity of the cerebral cortex in relation to musical instrument performance. Mark Jude Tramo (Dept. of Neurol., Harvard Med. School and Massachusetts General Hospital and the Inst. for Music and Brain Sci., Boston, MA 02114-2696, mtramo@hms.harvard.edu)

The acquisition and maintenance of fine-motor skills underlying musical instrument performance rely on the development, integration, and plasticity of neural systems localized within specific subregions of the cerebral cortex. Cortical representations of a motor sequence, such as a sequence of finger movements along the keys of a saxophone, take shape before the figure sequence occurs. The temporal pattern and spatial coordinates are computed by networks of neurons before and during the movements. When a finger sequence is practiced over and over, performance gets faster and more accurate, probably because cortical neurons generating the sequence increase in spatial extent, their electrical discharges become more synchronous, or both. By combining experimental methods such as single- and multi-neuron recordings, focal stimulation, microanatomical tracers, gross morphometry, evoked potentials, and functional imaging in humans and nonhuman primates, neuroscientists are gaining insights into the cortical physiology, anatomy, and plasticity of musical instrument performance.

9:45

5aMU3. Speed, accuracy, and stability of laryngeal movement in singing. Ingo R. Titze (Dept. of Speech Pathol. and Audiol., Univ. of Iowa, Iowa City, IA 52242 and Natl. Ctr. for Voice and Speech, Denver Ctr. for the Performing Arts, Denver, CO 80204, ititze@dcpa.org)

Motor performance is often quantified in terms of speed, strength, accuracy, and stability of a target gesture, or maintaining a given posture. In the vocal system, this involves primarily the intrinsic laryngeal muscles and the respiratory muscles. Agonist–antagonist pairs of muscles are used to position the vocal folds for phonation (vocal onset), for pitch change, and for registration (as in yodeling). Maximum speed and accuracy are discussed for vocal embellishments such as trills, trillo, scales, arpeggios, yodel, and glissando. This speed and accuracy are interpreted in terms of muscle twitch and tetanic responses obtained *in vitro* on animal muscles, from electromyographic recordings on humans, and from muscles not easily tested on humans. The laryngeal reflex system is also described, particularly with regard to its ability to stabilize (or destabilize) neurologic tremor originating from the central nervous system.

Contributed Paper

10:05

5aMU4. Flute "breath support" perception and its acoustical correlates. Isabelle A. Cossette and Patrick Sabourin (McGill Univ., Montreal, QC, Canada)

Music educators and performers commonly refer to ''breath support'' in flute playing, yet the term ''support'' is neither well-defined nor consistently used. Different breathing strategies used by professional flautists who were instructed to play with and without support were previously identified by the authors. In the current study, 14 musical excerpts with and without support were recorded by five professional flautists. Eleven professional flautists listened to the recordings in a random order and ranked (1 to 6) how much of the following sound qualities they judged to be in each example: support, intonation, control and musical expressiveness. Answers to the test showed that musical expressiveness was associated more closely with the supported excerpts than the answers about support itself. The ratings for each sound quality were highly intercorrelated. Acoustical parameters were analyzed (frequency and centroid variation within each note) and compared with the results of the perception test in order to better understand how the acoustical and psychological variables were related. The acoustical analysis of the central part of the notes did not show evident correlation with the answers of the perception test. [Work funded by the Social Sciences and Humanities Research Council of Canada.]

FRIDAY MORNING, 28 MAY 2004 CONFERENCE ROOM D, 7:50 A.M. TO 12:00 NOON

Session 5aNS

Noise: Noise in Large Cities I

Daniel R. Raichel, Chair
2727 Moore Lane, Fort Collins, Colorado 80526

Chair's Introduction—7:50

Invited Papers

8:00

5aNS1. Noise and soundscape in Rome. Giovanni Brambilla (CNR-Inst. of Acoust., via del Fosso del Cavaliere 100, 00133 Rome, Italy)

Noise pollution is an old problem in Rome. In 45 B.C. the Lex Julia Municipalis limited carriage traffic in the city center to specific times. Road traffic constitutes the most important and widespread noise source, and several investigations have been conducted since 1972, some aimed at developing a numerical model for predicting the hourly LAeq level. In order to reduce the large impact of this type of noise some measures have been carried out, including surfacing with porous asphalt, erection of noise barriers, limitation in time and spacing of private traffic, etc. However, most of the public complaints deal with noise from equipment operation and recreational activities rather than transportation systems. Moreover, the most famous tourist areas opened to pedestrians only are not as quiet as expected but their sound environment is usually rated more acceptable than noise from other sources at the same level. In compliance with the Italian legislation on noise, the Municipality of Rome issued a noise zoning code for its own territory, and a noise mapping is in progress, pursuant to the requirements of the 2002/49/EC European directive. A Geographical Information System has been also developed to manage all the aspects of noise pollution.

8:20

5aNS2. Environmental noise in Beijing. Jing Tian (Inst. of Acoust., Chinese Acad. of Sci., Beijing 100080, PROC) and Yi Wang (Beijing Ctr. of Environ. Protection and Monitoring, Beijing 100000, PROC)

Beijing is a city under rapid development. More than 60% of the complaints on the environment comes from noise pollution. With a road system of total length of 4200 km and thousands of construction sites, road traffic, and construction are the most important noise sources in Beijing. One third of its population of nearly 14 million inhabitants and more than 3 million visitors is influenced by traffic noise and 16% of its inhabitants are living along various roads and expressways. A series of noise control measures have been taken since the early 1980s, such as road system reconstruction and improvement, industrial noise source control, vehicle noise control, as well as various administrative steps, which reduced the traffic noise by 7 dBA with the number and average flux of vehicles multiplied by more than three times and kept the city in a day–night average level around 64 dBA. In this paper, the environmental noise legislation in Beijing is briefly introduced. The historical progress in noise control is reviewed. The noise pollution up-to-date is given and discussed. Some possible improvements and potential opportunities are presented.

8:40

5aNS3. The meaning of city noises: Investigating sound quality in Paris (France). Daniele Dubois, Catherine Guastavino, Valerie Maffiolo (CNRS LCPE/LAM 11 rue de Lourmel (F) 75015 Paris, France, ddubois@ccr.jussieu.frieu.fr), Catherine Guastavino (McGill Univ., Montreal, Canada), and Valerie Maffiolo (FTR&D, Lannion, France)

The sound quality of Paris (France) was investigated by using field inquiries in actual environments (open questionnaires) and using recordings under laboratory conditions (free-sorting tasks). Cognitive categories of soundscapes were inferred by means of psycholinguistic analyses of verbal data and of mathematical analyses of similarity judgments. Results show that auditory judgments mainly rely on source identification. The appraisal of urban noise therefore depends on the qualitative evaluation of noise sources. The salience of human sounds in public spaces has been demonstrated, in relation to pleasantness judgments: soundscapes with human presence tend to be perceived as more pleasant than soundscapes consisting solely of mechanical sounds. Furthermore, human sounds are qualitatively processed as indicators of human outdoor activities, such as open markets, pedestrian areas, and sidewalk cafe districts that reflect city life. In contrast, mechanical noises (mainly traffic noise) are commonly described in terms of physical properties (temporal structure, intensity) of a permanent background noise that also characterizes urban areas. This connotes considering both quantitative and qualitative descriptions to account for the diversity of cognitive interpretations of urban soundscapes, since subjective evaluations depend both on the meaning attributed to noise sources and on inherent properties of the acoustic signal.

9:00

5aNS4. The sound of Berlin. The noise annoyance you love to hate. Brigitte Schulte-Fortkamp and Cay Hehner (Inst. of Tech. Acoust., TU-Berlin, Einsteinufer 25, D-10587 Berlin, Germany, brigitte.schulte-fortkamp@tu-berlin.de)

Noise is part of life in Berlin. The Berlin Sound is often described as Berlin is a loud city, but this sound is also a matter of identification. Just like human beings, metropolitan areas may be identified by their gait. People enjoy living in Berlin but they also claim negative health effects and often permanent annoyance caused by daily environmental noise. For instance 1 119 714 cars demonstrate the volume of traffic from 6.30–9.00 a.m, and another impressive descriptor of the volume of traffic is the number of 1 567 600 cars from 4.00–6.30 p.m. In a representative survey of 2000 adults, almost all German (80%) citizens are affected by some level of noise pollution. The predominant source of noise in residential areas is road traffic, which remains a nuisance for over half the population, and a source of serious annoyance for some 18%. Next to road traffic, air traffic is the most important transport-related source of annoyance, followed by rail traffic noise. Results from a qualitative survey in a residential area of Berlin give insights into the different perspectives of the noise perception of the investigated subjects and of important parameters with respect to daily life.

9:20

5aNS5. Noise and social deprivation in an urban environment. Bridget Shield (Dept. of Eng., London South Bank Univ., London SE1 0AA, UK, shieldbm@lsbu.ac.uk) and Julie Dockrell (London Univ., London W1H 0AA, UK)

Noise levels have been measured inside and outside approximately 170 schools in London, England, as part of a project to investigate the effects of noise on the cognitive performance and academic attainments of children of primary school age in London. As well as providing data on individual schools the survey has provided a portrait of the noise climate across London. In addition to noise levels, the external noise sources present at each measurement location have been identified. The locations in which external schools noise levels were measured include areas where road traffic is the predominant noise source and areas near Heathrow Airport where aircraft are the major source. In addition to noise data, the following socioeconomic information has been obtained for a majority of the schools: percentages of children at each school having free school meals and the numbers for whom English is not the first language. The first of these is known to be a reliable indicator of social deprivation in an area. The relationship between noise and these socioeconomic factors has been established, which shows that, as might be expected, the higher noise levels in London are associated with the areas of greater social deprivation.

9:40

5aNS6. Noise in Mexico City. Sergio Beristain (IMA, ESIME, IPN. P.O. Box 12-1022, Narvarte, 03001, Mexico D. F., Mexico, sberista@hotmail.com)

Mexico City is known to be the largest city in the world, inhabited by some 20 percent of the national population, so noise pollution is not strange to it, particularly in view of the fact that industry is not concentrated, but rather spread throughout the city. The international airport also lies within the city limits, in the midst of residential areas. The heavy traffic during rush hours in the morning and in the evening and the activities of the populace, together with special events, produce a noise problem that is difficult to assess and to solve. Nevertheless, with educational programs begun several years ago and noise campaigns planned for the near future, in addition to existing regulations, the problem is not completely out of control. This paper presents a discussion of the general noise problem and describes how authorities and institutions are dealing with it.

10:00–10:20 Break

10:20

5aNS7. Noise in large cities in Brazil. Samir N. Y. Gerges (Mech. Eng. Dept. (EMC), Federal Univ. of Santa Catarina (UFSC), Florianopolis, SC, Brazil)

Large cities' noise is considered by the World Health Organization to be the third most hazardous pollution, preceded by air and water pollution. In urban centers, in general, and especially in developing countries such as Brazil, large populations are affected by excessive noise due mainly to traffic flow. The Brazilian Federal Government specifies noise limits, but each state can enforce its own set of noise limits, providing they are lower. The rapid economic growth, together with large migration of northern Brazilians to the

developing southern urban areas in search of more lucrative jobs in construction and industrial sectors, resulted in a fast increase in activities such as vehicle and bus traffic, home construction, and development of all necessary infrastructures to support this growth. Urban noise in Brazil has been receiving the attention of national authorities only since 1990, when the Federal Government approved the first ''Program of Community Silence,'' based on ISO R 1996-1971. This paper highlights the noise situation in the five largest and most populated cities in Brazil: Sao Paulo, Rio de Janeiro, Belo Horizonte, Porto Alegre and Curitiba [Zannin *et al.*, Appl. Acoust. **63**, 351–358 (2002)].

10:40

5aNS8. Noise issues in Kanagawa Prefecture. Shigenori Yokoshima (Kanagawa Environ. Res. Ctr. 1-3-39, Shinomiya, Hiratsuka, Kanagawa Prefecture 254-0014, Japan, yokoshima@k-erc.pRef. kanagawa.jp) and Akihiro Tamura (Yokohama Natl. Univ., Hodogaya, Yokohama 240-8501, Japan)

In Kanagawa Prefecture, bordering Tokyo Metropolis and the third most densely populated prefecture in Japan, various noises have caused serious problems in terms of living environment preservation and human health protection. This paper describes present states of noise issues in Kanagawa. Road traffic noise, remaining one of serious pollution issues, was monitored at a total 217 sites along trunk roads in Kanagawa from fiscal year 2000 to 2002. The percentage of the sites that achieve environmental quality standards for road traffic noise was approximately 20%. Noise caused by Tokaido Shinkansen trains, of which the total daily number is 287, also has negative impacts on inhabitants along the railway. As a result of the noise measurement from fiscal year 1994 to 2002, about 80% of the measurement sites exceeded environmental quality standards for Shinkansen railway noise during the years. In the areas surrounding the Atsugi Base, noise generated by training flights damagingly affects inhabitants' daily life. The number of complaints due to the noise was largest among noise issues. Moreover, neighborhood noises, noises emitted during the nighttime operation of bars, restaurants and shops, and noises produced by work in out-of-door yards have recently provoked social issues.

11:00

5aNS9. Managing environmental noise in Hong Kong. Kai Ming Li (Dept. of Mech. Eng., The Hong Kong Polytech. Univ., Hung Hom, Hong Kong)

Hong Kong is well known for its economic vibrancy and its hyper densely population: more than 7 million people living in a total area of slightly over 1000 square kilometers of hilly areas. Most of these people live and work in about 20% of the total land area, resulting in probably the highest densities in the world. The high population density is also matched by a large number of vehicles running in the roads. At present, there are over 400 000 vehicles operating on a highway network less than 1000 km in length. With all these factors plus many urban activities associated with the rapid growth and development, noise is an important environmental issue in the city. Although there are many dimensions for the quality of life, the acoustic environment is undoubtedly an essential part of it, especially when the people's aspirations increase as the society develops. This paper summarizes the development of strategies for controlling environmental noise in Hong Kong in the past two decades. The current situation will be addressed and a proposal for an improved traffic noise policy will be presented. [Work supported by the Research Grants Council, and The Hong Kong Polytechnic University.]

11:20

5aNS10. Noise enforcement in cities. Eric Zwerling (Rutgers Univ. Noise Tech. Assistance Ctr., Dept. of Environ. Sci., 14 College Farm Rd., New Brunswick, NJ 08901)

Noise enforcement programs (NEPs) in cities face a number of unique obstacles, including the necessity to conduct sound level measurements in a complex acoustical environment. It is important to regularly and objectively review the efficacy of the program in totality. A successful NEP consists of the following interactive components: the ordinance, trained enforcement personnel, sound measurement equipment, and prosecutorial mechanisms. If any of these components are not working in harmony, the NEP may not deliver the desired results. The NEP should be designed to reflect the realities of field enforcement, and should be tailored to the specific agency that will be conducting the enforcement, mindful of the differences between environmental compliance officers and police officers. This begins with the crafting of an ordinance that is appropriate for the unique conditions within the jurisdiction. The sound measurement equipment should be the least complicated necessary to take the specified measurements, and should be as durable as possible. Enforcement officers require proper training. Noncompliance must be addressed in a constructive and meaningful manner. A number of jurisdictions with active noise enforcement programs have been interviewed to determine the status of the program, and their strengths and weaknesses have been honestly assessed.

11:40

5aNS11. Noise impact evaluation method for supermarket sites. Gregory Tocci and Rose Mary Su (Cavanaugh Tocci Assoc., Inc., 327F Boston Post Rd., Sudbury, MA 01776, gtocci@cavtocci.com)

A large food supermarket chain is currently in a large store-building program involving several project managers who identify prospective sites, retain A/E design services, and obtain local permits for construction. There is a variety of environmental and planning issues that sometimes need special consideration; among these is environmental noise. Supermarket management has asked that consultants representing each discipline, and who work with project managers on specific projects as needed, issue guidelines as to when special consideration of their disciplines is required. This presentation describes a simple method developed for the supermarket owner that allows their project managers to judge whether there is a need for special consideration of environmental noise in store design by an acoustical consulting firm.

Session 5aPAa

Physical Acoustics: Wave Propagation: Homogeneous Media

Doug Meegan, Chair
Applied Research Laboratory, University of Austin, Austin, Texas 78713-8029

Contributed Papers

8:00

5aPAa1. Measurements and simulations of an asymmetric finite amplitude ultrasonic field. Kirk D. Wallace, Mark R. Holland, and James G. Miller (Dept. of Phys., Washington Univ. in St. Louis, 1 Brookings Dr., Campus Box 1105, St. Louis, MO 63130, kirk.wallace@wustl.edu)

Transducer arrays used routinely for medical imaging generate diffracting finite amplitude ultrasonic fields that are typically asymmetric and intricate. The goal of this work is to validate a Burgers equation enhanced, nonlinear angular spectrum simulation approach with experimental measurements of an asymmetric finite amplitude ultrasonic field. A one-dimensional transducer array with a 3:2 aspect ratio, separate azimuth and elevation foci, and a nonuniform aperture apodization was immersed in a watertank and was driven by a broadband pulse. A 0.6-mm-diam hydrophone receiver was mechanically scanned to obtain detailed maps of the harmonic content and spatial distribution of the finite amplitude ultrasonic field in nine transverse (00 to 160 mm, evenly spaced) and two orthogonal meridian plane cross sections. The fundamenal (2.3 MHz) component of the field, measured at face of the source, served as input for the numerical simulation (written in MATLAB and run on a notebook computer). Comparisons were performed at the fundamental, second, third, fourth, and fifth harmonics. Overall, excellent agreement was observed between experimental measurements of the nonlinear ultrasonic field generated with an asymmetric array transducer and the Burgers equation enhanced, nonlinear angular spectrum simulations. [Work supported by NIH R01 HL72761.]

8:15

5aPAa2. Numerical three-dimensional solution of broadband acoustic pulses through media with power-law attenuation. Margaret G. Wismer (Dept. of Elec. Eng., Bucknell Univ., Lewisburg, PA 17837, wismer@bucknell.edu)

Acoustic waves in tissues and weakly attenuative fluids often have an attenuation parameter, α, satisfying $\alpha=\alpha_0\omega^y$, in which ω is the applied frequency and y is between 1 and 2. This power-law attenuation is not predicted by the classical thermoviscous wave equation, and recent research has led to a number of modified viscous wave equations in which the third term usually consists of a convolution operator or a fractional spatial or temporal derivative. These wave equations are obtained by taking into account the requisite wave velocity dispersion predicted by the attenuation in order for the signals to be causal. In this paper, acoustic waves undergoing power-law attenuation are modeled by a slight modification to the thermoviscous wave equation, in which the time derivative of the viscous term is replaced by a fractional time derivative. This new equation satisfies the power-law formulation for lossy waves. An explicit time-domain, finite-element formulation leads to a stable algorithm capable of simulating three-dimensional broadband acoustic pulses propagating through attenuative and dispersive media. Results are given for pulse propagation through layered lossy media, and it is shown how attenuation affects the transmission and reflection of broadband signals.

8:30

5aPAa3. Numerical simulation of wave propagation in air including the effects of molecular relaxation and relative humidity. Mark S. Wochner, Anthony A. Atchley, and Victor W. Sparrow (Grad. Prog. in Acoust., Penn State Univ., University Park, PA 16802)

The research reported is directed at developing a computational model for assessing the importance of nonlinear propagation in the near field of jet engines. In order to attain accurate results, more complete models of acoustic wave propagation are required. The goal of the current research is to develop a two-dimensional time domain solution to the Navier–Stokes equations that includes the effects of classical absorption, relative humidity, and molecular relaxation of diatomic nitrogen and oxygen. Using fluid dynamics equations from Pierce [*Acoustics: An Introduction to Its Physical Principals and Applications* (ASA, New York, 1989)], a system of equations is created and solved using finite difference schemes. Benchmark cases will be discussed to show that the code is capable of reproducing the physical phenomena associated with atmospheric absorption and the relative importance of the various attenuation mechanisms will be discussed. [Work supported by ONR.]

8:45

5aPAa4. Numerical modeling of infrasound propagation at very long distance. Pierre-Franck Piserchia and Roger Roche (BP12, Bruyeres le chatel 91680, France, pierre-franck.piserchia@cea.fr)

Compliance with the CTBT in the atmosphere will be monitored by a world-wide network of infrasound stations consisting of 60 stations equipped with microbarographs in order to measure small changes in the air pressure in the frequency range 0.02 to 4 Hz. They are characterized by a good sensitivity, and by a large dynamic. By the application of array techniques, it is possible to determine the direction of pressure pulses caused by small explosions in the atmosphere, as well as shock waves caused by supersonic aircraft or meteorites. To take into account the nonlinear phenomena at the source and during the propagation, we are developing a numerical approach to solve the Euler nonlinear equation. In a first step, in the linear domain, this method is compared with two other numerical modeling approaches based on the ray tracing technique and the parabolic approach. In our test case, the source is on the ground and generates a 1-Pa pressure pulse centered at the frequency of 0.1 Hz. We considered an infrasound propagation over a distance of 500 km and an atmosphere height of 200 km. In a further step, the source level will be increased to study nonlinear phenomena.

9:00

5aPAa5. Terrain effects on acoustic pulse propagation. Donald G. Albert (USA ERDC-CRREL, 72 Lyme Rd., Hanover, NH 03755, Donald.G.Albert@erdc.usace.army.mil)

Outdoor sound propagation was measured at a variety of locations to determine terrain and environmental effects on acoustic pulses. Small explosive charges were detonated at a height of 1.5 m above the ground, and the resultant acoustic pulses were digitally recorded using pressure sensors located 10 to 100 m from the source. The experiments were conducted in an open area over concrete, grass, or snow; through a forest or thick tropical vegetation; and along a building-lined street in an artificial village. The measurements reveal that waveform changes occur at the shortest measurement ranges and increase as the propagation distance increases. These changes include amplitude attenuation, pulse width, and waveform shape and are caused by the differing ground conditions. The highest peak amplitudes at 100 m were recorded over concrete, and the lowest over snow-covered ground, with peak pressure decay rates of 24- and 32-dB/decade distance in these two situations. In addition, multiple scattering and reflected arrivals were visible in the tropical, forest, and urban areas. [Work funded by U.S. Army.]

9:15

5aPAa6. Modes and quasimodes for outdoor sound propagation at night. Roger Waxler (NCPA, Univ. of Mississippi, University, MS 38677, rwax@olemiss.edu)

Sound propagation outdoors at night is characterized by sound-speed profiles which are downward refracting in the first few hundred meters of the atmosphere but upward refracting at higher altitudes. The downward refraction causes sound to be ducted along the ground. Due to the upward refraction up high, however, this ducting is imperfect in that sound can leak away into the upper atmosphere. It has been shown that for most frequencies the sound is attenuated by the ground before it can leak out of the duct, effectively stabilizing the duct. For these frequencies a modal expansion has been developed which gives an efficient and physically transparent model for low-angle propagation. There are, however, certain resonant frequency bands in which the leaking can be significant. For these frequencies the modal expansion is not sufficient, but needs to be augmented by the addition of quasimodes. The quasimodes, while not strictly speaking modes, have the same form as the modes. A straightforward method for determining the quasimodes will be presented. The result is a modal expansion with uniform validity from frequencies of a few Hz up to a few kHz.

FRIDAY MORNING, 28 MAY 2004 ROYAL BALLROOM B, 9:45 A.M. TO 12:00 NOON

Session 5aPAb

Physical Acoustics: Wave Propagation: Time Reversal Methods

Evgenia A. Zabolotskaya, Chair
Department of Mechanical Engineering, University of Texas at Austin, 1 University Station, Austin, Texas 78712-0292

Contributed Papers

9:45

5aPAb1. Application of time reversal acoustics focusing for nonlinear imaging ms. Armen Sarvazyan and Alexander Sutin (Artann Labs., Lambertville, NJ, armen@artannlabs.com)

Time reversal acoustic (TRA) focusing of ultrasound appears to be an effective tool for nonlinear imaging in industrial and medical applications because of its ability to efficiently concentrate ultrasonic energy (close to diffraction limit) in heterogeneous media. In this study, we used two TRA systems to focus ultrasonic beams with different frequencies in coinciding focal points, thus causing the generation of ultrasonic waves with combination frequencies. Measurements of the intensity of these combination frequency waves provide information on the nonlinear parameter of medium in the focal region. Synchronized stirring of two TRA focused beams enables obtaining 3-D acoustic nonlinearity images of the object. Each of the TRA systems employed an aluminum resonator with piezotransducers glued to its facet. One of the free facets of each resonator was submerged into a water tank and served as a virtual phased array capable of ultrasound focusing and beam steering. To mimic a medium with spatially varying acoustical nonlinearity a simplest model such as a microbubble column in water was used. Microbubbles were generated by electrolysis of water using a needle electrode. An order of magnitude increase of the sum frequency component was observed when the ultrasound beams were focused in the area with bubbles.

10:00

5aPAb2. 3D shear wave generation in soft tissues using the time reversal kaleidoscopes. Delphine Palacio, Jeremy Bercoff, Gabriel Montaldo, Mickael Tanter, Mathias Fink (Laboratoire Ondes et Acoustique, 10 rue Vauquelin, 75005 Paris, France), Armen Sarvazyan, and Alexander Sutin (Artann Labs., Inc., Trenton, NJ)

The time reversal kaleidoscope provides elegant possibilities for both temporal and spatial concentration of acoustic energy in a wide frequency band. Using such ultrasound focused beams, it is possible to generate mechanical sources radiating low frequency shear waves inside tissues. Shear waves are of great interest to investigate tissue viscoelastic properties. In most techniques relying on this concept, generation and imaging of shear waves are realized with a single ultrasonic probe, limiting the capabilities concerning the generation of the shear source. Here we propose a system able to generate and image shear waves in a full 3D volume with any kind of polarization. The focused beam is generated by the time reversal kaleidoscope. By combining the concepts of time reversal mirrors and chaotic reverberating cavities, the time reversal kaleidoscope is able to focus ultrasound beams in a large 3D volume with only some tens of transducers and presents equivalent performances than conventional 2D matrices made of thousands of transducers. The kaleidoscope is triggered by an ultrafast imaging system able to image the resulting shear waves. *In vitro* results in tissue mimicking phantoms are presented. Chaotic cavities of different geometries are tested and compared. [Work supported in part by NIH grant.]

10:15

5aPAb3. Broadband time reversed acoustic focusing and steering system. Alexander Sutin (Artann Labs. and Stevens Inst. of Technol., NJ), Armen Sarvazyan (Artann Labs., Lambertville, NJ), Gabriel Montaldo, Delphine Palacio, Jeremy Bercoff, Mickael Tanter, and Mathias Fink (Laboratoire Ondes et Acoustique, 75005 Paris, France)

We present results of experimental testing and theoretical modeling of a time reversal acoustic (TRA) focusing system based on a multifaceted aluminum resonator with 15 piezoceramic transducers glued to the resonator facets. One of the facets of the resonator, a pentagon with characteristic dimension of about 30 mm, was submerged into a water tank and served as a virtual phased array which provided ultrasound focusing and beam steering in a wide frequency band (0.7–3 MHz). Ultrasonic pulses with different carrier frequencies and various complex waveforms were focused; the focal length was varied in the range of 10–55 mm and the focused beam was steered in a range of angles of ± 60 deg. The amplitude of the signal in the focal region reached 40 MPa. A theoretical model was based on an assumption that the radiating part of the resonator works as a phase conjugation screen for a spherical wave radiated from the focal point. Theoretical dependencies of the field structure on the position of the focus point and ultrasound frequency are in a good agreement with experimental results. TRA based focusing of ultrasound has numerous applications in medical diagnostics, surgery and therapy. [Work supported by NIH grant.]

10:30

5aPAb4. Acoustic tomographic array simulation. Sandra L. Collier, David A. Ligon, and John M. Noble (U.S. Army Res. Lab., 2800 Powder Mill Rd., Adelphi, MD 20783-1197)

Acoustic travel-time tomography of the atmosphere allows one to retrieve the virtual temperature and wind velocity fields, and to monitor their evolution in time. The temperature and wind velocity fields in a horizontal slice near the ground have been successfully retrieved by several research groups. One future direction for acoustic tomography of the atmosphere is the retrieval of detailed time-varying volumetric wind and temperature fields. The ability to retrieve these fields is strongly dependent upon the system configuration, e.g., the number and placement of sources and sensors, the geometry of the array, the environmental effects, and the type of the source. A sensitivity analysis for the retrieval of the temperature and wind velocity fields as a function of sensor configuration has been performed using an acoustic tomographic array simulator. We discuss the simulation model and results of the sensitivity analysis to some baseline propagation cases.

10:45

5aPAb5. Time reversal for source detection in urban environment. Lanbo Liu and Donald G. Albert (USA Cold Regions Res. and Eng. Lab., 72 Lyme Rd., Hanover, NH 03755-1290, lanbo.liu@erdc.usace.army.mil)

This paper applies the time reversal method to acoustic source detection in an outdoor urban environment. Experimental measurements were conducted in a full-scale artificial village to determine the effect of buildings on sound propagation outdoors. Explosive charges were detonated to produce acoustic pulses that were digitally recorded by sensors scattered throughout the village. We use the two-dimensional finite difference time domain (FDTD) method to compute synthetic time traces for acoustic propagation within the artificial village. The FDTD model predicted arrival times and amplitude levels that are in fair agreement with the measured data at most microphone measurement locations. Using the simulated time traces and the measured data from only nine stations and back propagating them into the model, the sound energy refocuses in the vicinity of the true source location. Our time reversal experiment confirms that the phase information is more critical than sound pressure levels and that using information acquired only at non-line-of-sight (NLOS) locations is sufficient to obtain accurate source locations. The results demonstrate that time reversal can potentially be used for fast source location in a complex urban terrain and noisy acoustic background, with without requiring LOS sensors. [Work funded by U.S. Army.]

11:00

5aPAb6. The time-reversal effect at sound scattering by a rough surface. Iosif Fuks and Konstantin Naugolnykh (Univ. of Colorado/Zeltech, Boulder, CO)

The scattering of the time-reversed wave by a rough surface is considered in the present paper. Using the Green function theory the point source field scattering by the rough surface in the Kirchhoff approximation is considered and the equation for the time-reversed and retransmitted back to the rough surface wave is obtained. The spatial distribution and the intensity spectrum of the scattered wave are evaluated for the Gaussian model of the rough surface and their features are examined. The modifications of the spectrum of the time-reversed wave as a function of the relation of the propagation time of the wave to the source-receiver array to the characteristic period of the surface roughness variation are discussed.

11:15

5aPAb7. Characterization of scattering object with the decomposition of the time-reversal operator: Theory and experiment. Jean-Gabriel Minonzio, Claire Prada (Laboratoire Ondes et Acoustique, CNRS, ESPCI, 10 rue Vauquelin, 75005 Paris France, claire.prada-julia@espci.fr), David H. Chambers (Lawrence Livermore Natl. Lab., Livermore, CA 94551), Dominique Clorennec, and Mathias Fink (CNRS, 75005 Paris, France)

The decomposition of the time-reversal operator provides information on the scattering medium. It has been shown [Chambers, J. Acoust. Soc. Am. **109**, (2001)] that a small spherical scatterer is in general associated with four eigenvalues and eigenvectors of the time-reversal operator. In this paper, the 2D problem of scattering by a cylinder measured by a linear array of transducers is considered. Experimental results are obtained for wires of different diameters and different materials. It is shown how the singular-value distribution and the singular vectors depend on the elastic wave speeds cL, cT, the density and the radius of each wire. These results offer a new perspective towards solution of the inverse problem by determining more than scattering contrast using conventional array processing like that used in medical ultrasonic imaging.

11:30

5aPAb8. Revisiting the Stokes relations in a time-reversal cavity: Suppression of intraplate echoes induced by a titanium plate. Jean-Francois Aubry, Francois Vignon, Michael Tanter, Gabriel Montaldo (Laboratoire Ondes et Acoustique, ESPCI, Université Paris VII, U.M.R. C.N.R.S. 7587, 10 rue Vauquelin, 75005 Paris, France, jf.aubry@espci.fr), and Mathias Fink (Université Paris VII, 75005 Paris, France)

When focusing through plates or tubes the presence of multiple interfaces induces reflected wavefronts that follow the main wavefront. Adaptive focusing techniques can be used to cancel the echoes. For that purpose, two linear arrays of transducers have been placed on each side of a titanium plate. Three propagation operators have been acquired: transmission from one array to the other, and two reflection operators acquired by each array. In this work, two adaptative focusing methods have been used to cancel the echoes: first, they have been suppressed with a time-reversal mirror, using the two arrays' cavity surrounding the plate. Second, the echoes have been canceled by using the inverse filter technique, inverting the transmission operator. Thus, the inverse filter achieves echo cancellation by using only the transmitted fields, whereas time reversal also requires the reflected fields. It is shown how transmission and reflection operators are related by the Stokes relations in a matrix formalism. These relations clearly exhibit how the inverse filter takes advantage of the reflections in the medium. An iterative mathematical resolution of these equations yields a new way to invert the transmission operator.

11:45

5aPAb9. Time reversal imaging of noise sources inside a reverberant room. Guillemette Ribay, Cédric Roux, Julien de Rosny, and Mathias Fink (Lab. Ondes et Acoustique, Université Denis Diderot, UMR CNRS 7587, ESPCI, 10 rue Vauquelin, 75005 Paris, France)

In a reverberant room, the multiple reflected echoes make imaging of acoustical random sources difficult. In this talk, we give an original solution based on Time Reversal Principles. Basically, in a first step, the acoustic field due to a noise source is recorded by a set of microphones. In a second step, all the recorded signals are time-reversed and re-emitted by loudspeakers that now replace the microphones. We observe a strong enhancement of the mean intensity around the initial source position. In a first part, we highlight the formal link between time-reversal of pulsed and noise sources in reverberant media. Especially we deduce that the signal to noise ratio depends only on the number of pairs of microphone/loud speaker for noise sources. Numerical simulations confirm these predictions. In a second part, we present experimental results performed in a $5*3*3=m$ and reverberant room. The intensity map of noise acoustic sources with a working bandpass of [100–3000] Hz has been performed. The same experiment has been carried out with two coherent or incoherent sources. Moreover, the frequency dependence of the resolution obtained with this technique is studied. Finally, we focus on the SNR with respect to the bandpass and central frequency.

FRIDAY MORNING, 28 MAY 2004 IMPERIAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 5aPPa

Psychological and Physiological Acoustics: Poster Session II

William P. Shofner, Chair
Parmly Hearing Institute, Loyola University of Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

5aPPa1. Spatial release from informational masking. Brad Rakerd (Dept. of Audiol. and Speech Sci., Michigan State Univ., East Lansing, MI 48824) and Neil L. Aaronson (Michigan State Univ., East Lansing, MI 48824)

A new method for investigating spatial release from informational masking was developed and employed in two experiments. The new method is computer controlled and efficient. It employs the versatile coordinate response measure speech stimulus set [Bolia *et al.*, J. Acoust. Soc. Am. **107**, 1065 (2000)]. The experiments were conducted in an anechoic room, with a primary loudspeaker in front of the listener and a secondary loudspeaker at 60 deg to the right. Target messages were presented from the primary speaker only. For a standard, distractor messages, simultaneous with the target, were also presented from the primary speaker only. Spatial release was measured by presenting the distractors from both primary and secondary speakers with a temporal offset. Experiment 1 fixed the offset (secondary leading, +4 ms) and varied the number of distractors (1 to 3) and the target-to-distractor ratio (−12 to +4 dB). Masking release, sometimes as large as 10 dB, was found for all combinations of these variables. Experiment 2 varied the offset over a wide range of values. Substantial release from masking was found for both positive and negative offsets, but only in the range in which speech echoes are suppressed (<50 ms). [Work supported by NIDCD grant DC 00181.]

5aPPa2. Localization of frequency-separated bands of noise in the median plane. Gerald Ng and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., 44 Cummington St. #427, Boston, MA 02215, geraldng@bu.edu)

Six subjects reported the perceived location of two sets of noise stimuli presented through speakers in the median-sagittal plane. Two sets of stimuli were generated by passing flat (1 to 16 kHz) Gaussian noise through multiple, equal-amplitude bandpass filters. The control stimulus included all frequencies between 1 and 16 kHz. The first set contained two to six equally spaced, 1/3-octave-wide noise bands. The second set contained five equally spaced noise bands that varied in common bandwidth. Loudspeakers in the medial-sagittal plane were located at six frontal locations (−30, −15, 0, 15, 30, 45, with 0 directly in front) and two rear locations (120, 180). The experiments were conducted in a sound-treated room with the speakers concealed from view. Stimuli were 200 ms in duration and presented from one speaker per trial. Subjects used a pen to mark perceived stimulus locations on preprinted coordinate diagrams. Despite noticeable intersubject differences, most subjects localized stimuli with five or six 1/3-octave-wide bands as well as control stimuli. Many subjects showed steady decreases in accuracy as the bandwidths of five bands were reduced. HRTFs were recorded for each listener, and several model algorithms are being evaluated. [Work supported by NIDCD: Grants R01 DC00100 and P30 DC04663.]

5aPPa3. Localization of multiple-band noises in the median plane. Gerald Ng and H. Steven Colburn (Hearing Res. Ctr., Boston Univ., 44 Cummington St. #427, Boston, MA 02215, geraldng@bu.edu)

Six subjects reported the perceived location of two sets of noise stimuli presented through speakers in the median-sagittal plane. Two sets of stimuli were generated by passing flat (1 to 16 kHz) Gaussian noise through multiple, equal-amplitude bandpass filters. The control stimulus included all frequencies between 1 and 16 kHz. The first set contained two to six equally-spaced, $\frac{1}{3}$-octave-wide noise bands. The second set contained five equally spaced noise bands that varied in common bandwidth. Loudspeakers in the medial-sagittal plane were located at six frontal locations (−30, −15, 0, 15, 30, 45, with 0 directly in front) and two rear locations (120, 180). The experiments were conducted in a sound-treated room with the speakers concealed from view. Stimuli were 200 ms in duration and presented from one speaker per trial. Subjects used a pen to mark perceived stimulus locations on preprinted coordinate diagrams. De-

spite noticeable intersubject differences, most subjects localized stimuli with five or six $\frac{1}{3}$-octave-wide bands as well as control stimuli. Many subjects showed steady decreases in accuracy as the bandwidths of five bands were reduced. HRTFs were recorded for each listener, and several model algorithms are being evaluated. [Work supported by NIDCD Grants R01 DC00100 and P30 DC04663.]

5aPPa4. Improved BTE hearing-aid directivity using a directional microphone array. Douglas L. Jones (Dept. of Elec. and Computer Eng., Univ. of Illinois at Urbana–Champaign, 1406 W. Green St., Urbana, IL 61801, dl-jones@uiuc.edu), Michael E. Lockwood (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801), Charissa R. Lansing (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801), and Albert S. Feng (Univ. of Illinois at Urbana-Champaign, Urbana, IL 61801)

Extraction of speech in noise is of great importance to hearing-impaired listeners. Directional microphones are incorporated in some hearing aids to improve noise rejection through increased directivity. A symmetric cardioid response is created either with a single directional microphone, or using beamforming with two omnidirectional microphones. The head-related transfer function (HRTF), however, introduces an asymmetry that cannot be exploited by a linear array of omnidirectional microphones. A new BTE array consisting of a gradient directional microphone with nulls in the front–back vertical plane and two omnidirectional microphones exploits the asymmetry of the HRTF to obtain almost 2 dB better directivity than the best cardioid. HRTFs measured on KEMAR with this array were transformed to the frequency domain, where directivity-maximizing coefficients in each band were derived. The Articulation-Index (AI) weighted directivity gain of this optimal three-element directional array was 6.4 dB greater than a single omnidirectional microphone on the BTE, whereas the directivity gain of the HRTF-optimized two-omni beamformer was 4.6 dB, and the optimal free-field cardioid placed on the head yielded 4.4 dB. [Work partially supported by NIH (NIDCD) under Grant No. 1 R01 DC005762-01A1.]

5aPPa5. Effect of source location and listener location on ILD cues in a reverberant room. Antje Ihlefeld and Barbara G. Shinn-Cunningham (Hearing Res. Ctr., Boston Univ., 677 Beacon St., Boston, MA 02215, ihlefeld@bu.edu)

Short-term interaural level differences (ILDs) were analyzed for simulations of the signals that would reach a listener in a reverberant room. White noise was convolved with manikin head-related impulse responses measured in a classroom to simulate different locations of the source relative to the manikin and different manikin positions in the room. The ILDs of the signals were computed within each third-octave band over a relatively short time window to investigate how reliably ILD cues encode source laterality. Overall, the mean of the ILD magnitude increases with lateral angle and decreases with distance, as expected. Increasing reverberation decreases the mean ILD magnitude and increases the variance of the short-term ILD, so that the spatial information carried by ILD cues is degraded by reverberation. These results suggest that the mean ILD is not a reliable cue for determining source laterality in a reverberant room. However, by taking into account both the mean and variance, the distribution of high-frequency short-term ILDs provides some spatial information. This analysis suggests that, in order to use ILDs to judge source direction in reverberant space, listeners must accumulate information about how the short-term ILD varies over time. [Work supported by NIDCD and AFOSR.]

5aPPa6. Itakura distance measure used to study speech intelligibility in a room using white noise. Jorge Sommerhoff, Jorge Cardenas, Victor Poblete, and Jose Luis Barros (Instituto de Acustica, Universidad Austral de Chile, Campus Miraflores, Valdivia, Chile, jsommerh@uach.cl)

The recognition process of isolated words use distance measures to determine if the word corresponds to one of its vocabulary. In this investigation Itakura distance measure was used to compare the segments of white noise of two signals with different relationships of direct and reverberant energy densities. Interesting data is obtained from the distance measure in the function of the position of the microphone in the room in which a white noise source is continuously working. These results could be useful to develop a new methodology to measure the word intelligibility in a room. There were carried out some comparative measurements with Rasti and intelligibility test. [Work supported by DID–UACh.]

5aPPa7. Perceptual scaling of room reverberation. Pavel Zahorik (Dept. of Psychol. and Brain Sci., Univ. of Louisville, Louisville, KY 40292, pavel.zahorik@louisville.edu)

Recent evidence suggests that reverberant energy can provide listeners with important spatial information regarding the distance of a sound source. However, relatively little is known about the perceptual attributes of the reverberation itself, and how these attributes may be related to physical properties of the environment that also potentially impact perceived spatial location. Here, perceived similarity among 15 reverberant rooms simulated using virtual auditory space techniques was examined. Room size and surface absorption properties were varied, along with aspects of the virtual simulation including the use of individualized head-related transfer function (HRTF) measurements and properties of the room acoustic simulation. Seven listeners rated perceived similarity on a 100-point scale between all possible pairs of simulated rooms using a speech source signal. Multidimensional scaling techniques were used to estimate scales of perceived room reverberation. Although the resulting scales were complex and somewhat unique to individual listeners, it is clear that the perceptual effects of manipulating properties of the reverberant sound are much larger than the effects due to either nonindividualized HRTFs or nonoptimal room simulation methods. [Work supported by NIDCD.]

5aPPa8. Disrupting the build-up of the precedence effect. Richard L. Freyman (Dept. of Commun. Disord., Univ. of Massachusetts, Amherst, MA 01003, RLF@comdis.umass.edu) and Rachel Keen (Univ. of Massachusetts, Amherst, MA)

When pairs of clicks are presented with a brief delay to one loudspeaker, the perception of a single fused image builds up as the clicks are repeated. If lead and lag loudspeakers are then reversed, the precedence effect resets, i.e., two auditory images are often heard immediately after the switch. However, Djelani and Blauert [Acta Acust. (Beijing) **87**, 253–261 (2001)] found that a brief reversal in lag location did not reset the built-up fusion at the original spatial configuration. The current experiment determined whether the build-up effect would be disrupted or slowed by interleaving variations in lead–lag stimuli within a train of repeated clicks. During the standard train the lead was presented from 45 deg left and the lag from 45 deg right, the click pairs were presented at a rate of 4/s, the number of clicks ranged from 3 to 9, and the delay between lead and lag varied from 2 to 14 ms across trials. Results show that the variation that disrupted the build-up most consistently was when clicks from only the lead loudspeaker were interleaved. This condition simulated the absence of expected reflections. [Work supported by NIDCD DC01625.]

5aPPa9. A real-time virtual auditory system for spatially dynamic perception research. Jacob W. Scarpaci and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed Eng., Boston Univ., 44 Cummington St., Boston, MA 02215, scarpaci@bu.edu)

A Real Time Virtual Auditory System (RT-VAS) is being developed to provide a high-performance, cost-effective, flexible system that can dynamically update filter coefficients in hard real time on a PC. An InterSense head tracker is incorporated to provide low-latency head tracking to allow studies with head motion. Processing is done using a real-time Linux Kernel (RTAI kernel patch) which allows for precise processor scheduling, resulting in negligible time jitter of output samples. Output is calculated at a sample rate of 44.1 kHz and displayed using a National Instruments DAQ. Object oriented approach to system development allows for customizable input, position, and calculation routines as well as multiple independent auditory objects. Input and position may be calculated in real-time or read from a file. Calculation of output may include filtering with spatially sampled HRTFs or analytic models and head movements may be recorded to file. Limitations of the system are tied to the speed of the processor, thus complexity of experiments scales with speed of computer hardware. The current system handles multiple moving sources while tracking head position. Preliminary psychoacoustic results with head motion will be shown, as well as a demonstration of the system. [Work supported by NIH DC00100.]

5aPP10. Temporal integration of interaural-time differences carried by trains of 500-Hz tone bursts. Michael A. Akeroyd (MRC Inst. of Hearing Res., Glasgow Royal Infirmary, 16 Alexandra Parade, Glasgow G31 2ER, UK, maa@ihr.gla.ac.uk)

The just-noticeable difference for the interaural time difference (ITD) of a sound generally reduces as its duration is increased [e.g., T. Houtgast and R. Plomp, J. Acoust. Soc. Am. **44**, 807–812 (1968); E. Hafter and R. H. Dye, J. Acoust. Soc. Am. **73**, 644–651 (1983)]. The rate of this temporal integration is less than would be expected from a simple multiple-looks application of signal detection theory, as though one part of the sound is given greater weighting. The present experiment measured the contributions of the start, middle, or end of a sound to the detectability of its ITD. The stimulus was a 32-pip train of 500-Hz, 10-ms pips, diotic apart from some (2–16) target pips which carried the ITD to be detected and which were placed either at the start, middle, or end of the train. Integration rates were determined from psychometric functions measured using four normal-hearing listeners. The results showed that the rate was least for targets placed at the start or end but larger for targets at the middle. The data can be described using a weighted multiple-looks approach, based on an integration function of approximately 100-ms time constant together with an emphasis of the start, and end, of the sound.

5aPPa11. A shot-noise phase-opponency model for monaural detection of pure tones in noise at low frequencies. Yan Gai and Laurel H. Carney (Inst. for Sensory Res. & Dept. of Bioengineering & Neurosci., Syracuse Univ., Syracuse, NY 13244, mountfall@yahoo.com)

Phase-locked responses of low-frequency auditory-nerve (AN) fibers contain sufficient information for detection of a tone in noise. The phase-opponency (PO) model [Carney *et al.*, Acustica **88**, 334–347 (2002)] uses relative phase across fibers with different characteristic frequencies (CFs). The model's discharge rate is reduced when a tone is added to noise due to changes in the coincidence of the rate functions across AN CFs. In the current study, a shot-noise model for coincidence detection was used to implement the PO model. AN discharge times were simulated using an AN model with a discharge generator; therefore, internal noise was naturally included in the AN input fibers. Model results were compared to human data for detection of a 500-Hz tone in reproducible noises, which provided external noise. The aim was to test the hypothesis that the shot-noise PO model explains human performance in terms of thresholds, hit- and false-alarm rates, and to understand the performance for different reproducible noises. The ratio of the internal to external noise in the shot-noise PO model depended on combining responses of several identical but independent cells. [Work supported by NIDCD R01-01641.]

5aPPa12. Comparing spectral resolvability in chinchillas and human listeners using phase discrimination. William P. Shofner, Kathryn Sparks, Yuanxing Esther Wu, and Ellen Pham (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, wshofne@luc.edu)

A tone complex made of harmonic components that are added in cosine-starting phase can be discriminated from complexes comprised of identical harmonics that are added with random-starting phases. Phase discrimination occurs when unresolved harmonics interact within a single auditory channel. When harmonics are resolved, there is less interaction among components resulting in poorer phase discrimination performance. Thus, phase discrimination indirectly reflects spectral resolvability. Performance in a phase discrimination task was measured in chinchillas and human listeners to compare spectral resolvability between the two groups. Subjects discriminated a cosine-phase tone complex from random-phase tone complexes in a go/no-go behavioral paradigm. Tone complexes were comprised of a 250-Hz fundamental frequency and N consecutive higher harmonics, where N was 5, 10, 20, and 40. Performance was evaluated in terms of d'. The results show that the measured d' increased as N increased, and values of d' for each N condition were similar between chinchillas and human listeners. Values of the criterion for each N condition were also similar between chinchillas and humans. The results do not support the hypothesis that spectral resolvability is poorer in chinchillas, but suggest that resolvability is similar between the two groups. [Work supported by NIH/NIDCD.]

5aPPa13. Mechanisms of forward masking. Magdalena Wojtczak and Neal F. Viemeister (Dept. of Psych., Univ. of Minnesota, 75 East River Rd., Minneapolis, MN 55455, wojtc001@umn.edu)

Experiments involving forward masking were run with the aim of elucidating the mechanism underlying this phenomenon. Two possible mechanisms were considered: adaptation following masker offset, and persistence of excitation. In one experiment, detection of a 10-ms noise probe following a 150-ms noise masker was measured as a function of the level of a 20-ms burst of noise temporally centered on the probe. Model predictions were generated assuming either a reduction of gain applied to the probe (adaptation) or additivity of persisting masker excitation and excitation produced by the probe. Both models accounted for the data equally well. In the second experiment, listeners performed an image-centering task involving binaural presentation of a 4-kHz 10-ms probe. In one ear, the probe was preceded by a 150-ms noise forward masker. The level of the probe in the unmasked ear was adjusted until the image of the probe was centered. In general, the image was centered when the probe level in the unmasked ear was much lower (by 10–20 dB) than that in the masked ear. This result could be explained in terms of adaptation underlying forward masking but not in terms of persistence of excitation. [Work supported by NIH/NIDCD DC00683.]

5aPPa14. Discrimination of common-envelope signals: Comparison of static and dynamic frequency sounds. Lawrence Feth, Dean Hudson (Dept. of Speech and Hearing Sci., Ohio State Univ., Columbus, OH 43210, feth.1@osu.edu), Daniel Hack, Vivek Rajendran, and Ashok Krishnamurthy (Ohio State Univ., Columbus, OH 43210)

Listeners can match the rate of frequency change in an FM glide to the transition rate heard in a virtual frequency (VF) glide. A VF glide is perceived in a two-tone complex when the amplitude of one tone increases, while that of the other tone decreases, over the duration of the signal. This phenomenon is robust for signals spanning five or more au-

ditory filter bands (ERBs). Static two-tone complexes that approximate the end points of the VF glide are discriminable for components separated by more than 20 Hz, but less than five ERB. Preliminary work on discrimination of VF glides versus FM glides processed to have the same envelope as the VF signal led to estimates of just discriminable bandwidth less than one ERB (Rajendran *et al.*, 2003). Center frequencies ranged from 500 to 4000 Hz, but duration was fixed at 250 ms. The preliminary study has been extended to include a range of signal durations from 25 to 250 ms, and to determine the maximum resolution bandwidth for each combination of duration and center frequency. Predictions of the perceptual spectral centroid model will be compared to the results from well-practiced listeners.

5aPPa15. Scaling the perceived fluctuation strength of frequency-modulated tones. Florian Wickelmaier and Wolfgang Ellermeier (Dept. of Acoust., Aalborg Univ., Fredrik Bajers Vej 7 B5, 9220 Aalborg East, Denmark, fw@acoustics.dk)

Fluctuation strength is one of the major psychoacoustic variables considered in sound-quality evaluation. Zwicker and Fastl [*Psychoacoustics* (Springer, Berlin, 1999)] summarize recommendations for its computation, which have already been implemented in various software applications, even though the data basis is rather limited. In particular, the dependency of fluctuation strength on modulation frequency and modulation depth has seemingly never been tested in a factorial design. Therefore, in experiment I both of these factors were varied simultaneously in order to create 54 different frequency-modulated sinusoids. The task of the subjects was to directly estimate the perceived magnitude of fluctuation strength. The results do not conform well with the prevalent model of fluctuation strength. In experiment II this finding was further investigated by varying only one factor at a time. The results show that large individual differences, particularly in the effect of modulation frequency, persist. Thus, in experiment III the interaction of both factors was analyzed on an individual basis. By employing a 2AFC procedure, matches in fluctuation strength were obtained. The results suggest that most listeners are not able to integrate modulation frequency and modulation depth additively into a unidimensional percept. [Work supported by Centercontract on Sound Quality, Aalborg University.]

5aPPa16. Effect of noise on the detection of intensity increments. Walt Jesteadt, Lance Nizami, Stephen T. Neely, and Kim S. Schairer (Boys Town Natl. Res. Hospital, 555 N. 30th St., Omaha, NE 68131, jesteadt@boystown.org)

The detection of an increment in a longer duration tone is often used as a measure of intensity resolution. Many studies of increment detection have used background noise, but few have varied the level of the longer duration tone relative to the noise or have compared results obtained with or without the noise. To explore this space in more detail, thresholds were obtained by adaptive tracking from four subjects for detection of a 50-ms increment in a 350-ms, 4000-Hz tone presented in quiet or in broadband noise of 0, 10 and 20 dB spectrum level. Tone levels ranged from 15 to 75 dB SPL in quiet and from approximately 0 to 30 dB above the tone's detection threshold in the noise. Increment detection thresholds expressed as values of the difference in overall level of the tone resulting from addition of the increment increase as tone sensation levels decrease and are greater in noise than in quiet, but increase only slightly as a function of noise level. The data are well fitted by a model that assumes a specific form of nonlinearity motivated by loudness-matching data. [Work supported by NIDCD.]

5aPPa17. Loudness recalibration at short ISI: A closer look. Yoav Arieh (Dept. of Psych., Montclair State Univ., Montclair, NJ 07043, ariehy@mail.montclair.edu), Jennifer R. Mailloux (Mary Washington College, Fredericksburg, VA 22401-5358), and Lawrence E. Marks (John B. Pierce Lab., New Haven, CT 06519)

The loudness of a moderate level tone is substantially reduced when preceded by a louder tone of the same frequency (loudness recalibration). The reduction in loudness depends on the interstimulus interval (ISI) between the tones. Arieh and Marks (2003) reported that loudness starts to decline only when ISI exceeds about 200 m. One possible explanation is that loudness reduction takes 200 ms to develop. Alternatively, the delay in loudness reduction could reflect the combination of two processes: short-term facilitation and longer-term suppression. To test this hypothesis, subjects compared the loudness of a roving-level 500-Hz tone to a 60-dB 2500-Hz target tone presented, first, 100 ms after an 80-dB 2500-Hz recalibration tone and, subsequently, with the recalibration tone omitted; by omitting the recent recalibration tone, we hoped to observe the long-term effects on loudness of earlier recalibration tones. As expected, the loudness of the target was unaffected when it followed the recalibration tone by 100 ms, but loudness declined significantly when the recalibration tone was subsequently omitted. These results are consistent with the hypothesis that at short ISI, a recalibration tone sets off long-term suppression of loudness that is initially offset by short-term facilitation.

5aPPa18. Simple reaction time to narrow-band and broadband noise. Eva Wagner, Mary Florentine (Inst. of Hearing, Speech & Lang. and SLPA Dept. (106A FR), Northeastern Univ., 360 Huntington Ave, Boston, MA 02115, e.wagner@neu.edu), Søren Buus, and Joseph McCormack (Northeastern Univ., Boston, MA 02115)

The present study evaluates the relationship between loudness matches and simple reaction times (RT) in six normal-hearing subjects. Loudness matches between a narrow-band noise (125 Hz wide) and a broadband noise (1500 Hz) centered at 1000 Hz were made at levels from near threshold to near 100 dB SPL. For the same noises and level range, RT were also measured. In agreement with previous loudness-matching studies, as SPL increased the level difference between the noises needed to maintain equal loudness first increased, to around 10 dB at moderate SPLs, and then decreased. Except for one listener, the RT data show the same pattern. The level difference needed to maintain RT the same to the two noises first increased and then decreased. These results show that RT is closely related to loudness, but not to sensation level. If RT depended on sensation level, the level difference between the two noises needed to achieve equal RT would not change with SPL because the difference in sensation level between two sounds is constant as a function of SPL. Overall, the present data provide strong support for the contention that simple RT depends strongly on loudness. [Supported by NIH/NIDCD Grant No. R01 DC 02241.]

5aPPa19. Relative intelligibility of dynamically extracted transient versus steady-state components of speech. J. R. Boston, Sungyub Yoo, C. C. Li, Amro El-Jaroudi (Dept. of Elec. Eng., Univ. of Pittsburgh, Pittsburgh, PA 15261, boston@engr.pitt.edu), J. D. Durrant, Kristie Kovacyk, and Stacey Karn (Univ. of Pittsburgh, Pittsburgh, PA 15261)

Consonants are recognized to dominate higher frequencies of the speech spectrum and to carry more information than vowels, but both demonstrate quasi-steady state and transient components, such as vowel to consonant transitions. Fixed filters somewhat separate these effects, but probably not optimally, given diverse words, speakers, and situations. To enhance the transient characteristics of speech, this study used time-

varying adaptive filters [Rao and Kumaresan, IEEE Trans. Speech Audio Process. **8**, 240–254 (2000)], following high-pass filtering at 700 Hz (well-known to have minimal effect on intelligibility), to extract predominantly steady-state components of speech material (CVC words, NU-6). The transient component was the difference between the sum of the filter outputs and the original signal. Psychometric functions were determined in five subjects with and without background noise and fitted by ogives. The transient components averaged filtered speech energy, but PBmax was not significantly different (nonparametric ANOVA) from that of either the original or highpass filtered speech. The steady-state components yielded significantly lower PBmax (p 3D 0.003) despite their much greater energy, as expected. These results suggest a potential approach to dynamic enhancement of speech intelligibility. [Work supported by ONR.]

5aPPa20. Contribution of consonant versus vowel information to sentence intelligibility by normal and hearing-impaired listeners. T. Zachary Burkle, Diane Kewley-Port, Larry Humes, and Jae Hee Lee (Speech and Hearing Sci., Indiana Univ., Bloomington, IN 47405, tzachbur@indiana.edu)

The purpose of this study was to examine the contribution of information provided by vowels versus consonants to sentence intelligibility in young normal-hearing (YNH) and elderly hearing-impaired (EHI) listeners. Sentences were presented in three conditions, with either the vowels or the consonants replaced with speech shaped noise, or unaltered. Sentences from male and female talkers in the TIMIT database were selected. EHI subjects listened at 95 dB SPL, and YNH subjects at both 95 and 70 dB SPL. Subjects listened to each sentence twice and were asked to repeat the entire sentence after each presentation. Words were scored correct if identified exactly. Average performance for unaltered sentences was greater than 94%. Vowel-present conditions were always significantly more intelligible than consonant-present conditions, similar to data reported by Cole and colleagues [Proceedings of ICASSP, 1996]. Across groups, performance in the vowel-present conditions exceeded that in the consonant-present conditions by 14% to 40%, although EHI subjects performed more poorly than YNH subjects. In contrast to written English, vowels in spoken language carry more information about sentences than consonants for both normal and hearing-impaired listeners. [Work supported by NIDCD-02229.]

5aPPa21. Speech reception thresholds in various interference conditions. Suzanne P. Carr and H. Steven Colburn (Hearing Res. Ctr. and Dept. of Biomed Eng., Boston Univ., Boston, MA 02215)

Speech intelligibility is integral to human verbal communication; however, our understanding of the effects of competing noise, room reverberation, and frequency range restriction is incomplete. Using virtual stimuli, the dependence of intelligibility threshold levels on the extent of room reverberation, the relative locations of speech target and masking noise, and the available frequency content of the speech and the masking noise is explored. Speech-shaped masking noise and target sentences have three spectral conditions: wideband, high pass above 2-kHz, and low pass below 2-kHz. The 2-kHz cutoff was chosen to approximately bisect the range of frequencies most important in speech, and the high pass noise condition simulates high-frequency hearing loss. Reverberation conditions include a pseudo-anechoic case, a moderately reverberant ''classroom'' case, and a very reverberant ''bathroom'' case. Both binaural and monaural intelligibility are measured. Preliminary results show that source separation decreases thresholds, reverberation increases thresholds, and low frequency noise reverberates more in the rooms, contributing to increasing thresholds along with the effects of the upward spread of masking. The energetic effects of reverberation are explored. [Work supported by NIH DC00100.]

5aPPa22. Revised audibility-noise variance of speech recognition sensitivity model. Hannes Müsch (Sound ID, 3430 W. Bayshore Rd., Palo Alto, CA 94303, hmuesch@soundid.com) and Søren Buus (Northeastern Univ., Boston, MA 02115)

The speech recognition sensitivity (SRS) model [H. Müsch and S. Buus, J. Acoust. Soc. Am. **109**, 2896–2909 (2001)] was used to predict normal-hearing listeners' consonant-recognition performance in a nonsense-syllable context as measured in nine independent studies. By attempting to fit these data, the relation between the variance of the SRS model's audibility noise σ_N^2 and the speech-excitation to noise-excitation ratio SNR_E in the auditory periphery was optimized. The best predictions were obtained when the audibility-noise variance was directly proportional to the noise-excitation power. Based on this finding it is suggested that the relation between audibility-noise variance and SNR_E that was proposed in the original SRS paper be replaced by $10\log_{10}(\sigma_{N_i}^2) = -SNR_{E_i} + B + 10\log_{10}((z_j - z_i)/(19\text{Barks}))$, where $\sigma_{N_i}^2$ is the audibility-noise variance in the band reaching from critical-band rate z_i to z_j, SNR_{E_i} is the speech-excitation to noise-excitation ratio in that band expressed in dB, and B is a fitting constant that will be in the order of 13 dB. The model provided good predictions of the data with model parameters that are consistent with those of earlier studies. [Work supported by NIH/NIDCD Grant R01DC00187.]

5aPPa23. Lateral suppression preserves speech intelligibility at high intensities. James A. Bashford, Jr., Richard M. Warren, and Peter W. Lenz (Dept. of Psych., Univ. of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, bashford@uwm.edu)

The intelligibility of narrowband speech decreased 23% when its level was raised from 45 to 75 dB. However, when flanking bands of low-pass and high-pass white noise were added, intelligibility at the higher speech level recovered by as much as 65%. Recovery appears to be due to lateral suppression counteracting overloading effects of auditory-nerve (AN) firing-rate saturation at high speech intensities. Findings supporting this hypothesis include: (1) the absence of intelligibility enhancement at the lower speech level; (2) a greater effect of the higher frequency flanking noise band at a low noise level; and (3) equivalent effects of continuous versus gated flanking noise, inconsistent with firing-rate adaptation. In addition, there was no intelligibility increase when the flanking noise and narrow-band speech were delivered to opposite ears of listeners. These behavioral results are consistent with previous physiological observations and models of lateral suppressive interactions occurring in the lower auditory pathway. It appears that the noise-induced intelligibility recovery is produced via lateral inhibition of saturated AN-fiber input to neurons of the cochlear nucleus, with a possible additional contribution from mechanical (two-tone) suppression evoked within the cochlea and likely not involving olivocochlear feedback. [Work supported by NIH.]

5aPPa24. Development of a topic-related sentence corpus for speech perception research. Karen S. Helfer and Richard L. Freyman (Univ. of Massachusetts, Amherst, MA 01003)

A large sentence corpus has been developed for use in speech recognition research. Sentences ($n = 881$, three scoring words per sentence) were developed under 23 topics. In the first phase of development subjects rated each individual scoring word for relatedness to its given topic on a Likert scale. Next, two groups of young, normal-hearing listeners ($n = 16$/group) listened and responded to the recordings of the sentences (spoken by a female talker) presented with one of two types of maskers: steady-state noise ($S:N = -13$ dB) or two other females speaking random sentences ($S:N = -8$ dB). Each subject responded to half of the sentences with topic supplied and half with no topic supplied. Data analyses focused on addressing two questions: whether supplementation of topic would be more important in the presence of the speech masker versus the noise masker, and how the degree of relatedness of each key word to the topic influenced the effect of topic on recognition. The data showed little dif-

ference in how beneficial the topic was for speech versus noise maskers. Moreover, there was a complex relationship between effect of topic, type of masker, and position of the word in the sentence. [Work supported by NIDCD DC01625.]

5aPPa25. Immersive simulation of hearing loss and auditory prostheses. Patrick M. Zurek and Joseph G. Desloge (Sensimetrics Corp., 48 Grove St., Somerville, MA 02144, pat@sens.com)

Simulation of hearing loss is useful for demonstrating the communication challenges facing hearing-impaired people. However, current simulations, most of which are only recordings, do not actually elevate thresholds; i.e., they do not simulate hearing loss, *per se*. The hearing loss simulator described in this talk is immersive; the user's detection thresholds for ambient sounds are shifted by a prescribed degree. This threshold shift is achieved through a combination of passive attenuation (from muff-type hearing protectors) and additive masking noise (introduced by within-muff earphones). Acoustic signals picked up by microphones near each ear are processed through bandpass AGC channels and delivered via the earphones to complete the simulation of frequency-dependent hearing loss and loudness recruitment. Preliminary results validating the accuracy of specified threshold shift will be presented, along with speech-reception data comparing simulated with actual hearing losses. Subjective reactions of users engaged in one-on-one conversation suggest that strong feelings of communication disability are engendered by even moderate degrees of simulated hearing loss. The system, which is capable of simulating any degree of recruiting hearing loss along with hearing aids or cochlear implants, can provide effective interactive demonstrations of both auditory communication handicap and rehabilitation options. [Work supported by NIDCD.]

FRIDAY MORNING, 28 MAY 2004 CONFERENCE ROOM E, 8:45 TO 11:15 A.M.

Session 5aPPb

Psychological and Physiological Acoustics: More Complex Stimuli

Elizabeth A. Strickland, Cochair
Audiology and Speech Sciences, Purdue University, 1353 Heavilon Hall, West Lafayette, Indiana 47907

Jungmee Lee, Cochair
Speech and Hearing Sciences, City University of New York Graduate Center, 365 Fifth Avenue, New York, New York 10016

Contributed Papers

8:45

5aPPb1. Uncertainty and confusion in temporal masking. C. Formby (Div. of Otolaryngol.-HNS, Univ. of Maryland School of Medicine, 16 S. Eutaw St., Ste. 500, Baltimore, MD 21201, cformby@smail.umaryland.edu) and T. Zhang (Univ. of Maryland, College Park, MD 20742)

In a landmark study, Wright *et al.* [Nature **387**, 176–178 (1997)] reported an apparent backward-masking deficit in language-impaired children. Subsequently, these controversial results have been influential in guiding treatments for childhood language problems. In this study we revisited Wright *et al.*'s temporal-masking paradigm to evaluate listener uncertainty effects. Masked detection was measured for 20-ms sinusoids (480, 1000, or 1680 Hz) presented at temporal positions before, during, or after a gated narrowband ($W = 600–1400$ Hz) masker. Listener uncertainty was investigated by cueing various stimulus temporal properties with a 6000-Hz sinusoid presented either ipsi- or contra-lateral to the test ear or bilaterally. The primary cueing effect was measured in the backward-masking condition for a contralateral cue gated simultaneously with the on-frequency 1000-Hz signal. The resulting cued masked-detection threshold was reduced to quiet threshold. No significant cueing effects were obtained for other signal temporal positions in the masker nor for any off-frequency signal conditions. These results indicate that (1) uncertainty can be reduced or eliminated for on-frequency backward masking by cueing the signal and (2) the deficit reported by Wright *et al.* for language-impaired children may reflect uncertainty and confusion rather than a temporal-processing deficit *per se*. [Research supported by NIDCD.]

9:00

5aPPb2. Effect of masker onset asynchrony on overshoot in simultaneous masking. Andrzej Miskiewicz (Dept. of Speech-Lang. Pathol. and Audiol. (106A FR), Inst. of Hearing, Speech and Lang., Northeastern Univ., Boston, MA 02115), Soren Buus, and Mary Florentine (Northeastern Univ., Boston, MA 02115)

This study examines how overshoot is influenced by asynchrony between the onsets of an on-frequency center-band noise masker (CB) and a flanking noise band (fringe). Thresholds were measured for 2-ms tones at 5 kHz gated on 2 ms after the onset of the CB. Measurements were made with only the CB (4590–5464 Hz), and with either a low fringe (1900–4590 Hz) or a high fringe (5500–11 000 Hz) added to the CB. The fringe came on between 500 ms before and 100 ms after the CB onset. Because previous studies demonstrated that overshoot is produced only by masker components outside the auditory filter centered at the signal frequency, overshoot is taken as the increase in threshold caused by adding the fringe. Results show that considerable overshoot, 7–11 dB, is produced by a high fringe gated on within 10 ms before or after CB onset; overshoot is maximum when the fringe and CB come on simultaneously. Maximum overshoot with a low fringe is less than 4 dB. The finding that a high fringe increases threshold even when it comes on shortly after the signals end suggests that overshoot is governed primarily by central mechanisms. [Work supported by NIH/NIDCD Grant R01DC00187.]

9:15

5aPPb3. Audibility of time-varying signals in time-varying backgrounds: Model and data. Brian C. J. Moore and Brian R. Glasberg (Dept. of Exp. Psych., Univ. of Cambridge, Downing St., Cambridge CB2 3EB, England, bcjm@cus.cam.ac.uk)

We have described a model for calculating the partial loudness of a steady signal in the presence of a steady background sound [Moore *et al.*, J. Audio Eng. Soc. **45**, 224–240 (1997)]. We have also described a model for calculating the loudness of time-varying signals [B. R. Glasberg and B. C. J. Moore, J. Audio Eng. Soc. **50**, 331–342 (2002)]. These two models have been combined to allow calculation of the partial loudness of a time-varying signal in the presence of a time-varying background. To evaluate the model, psychometric functions for the detection of a variety of time-varying signals (e.g., telephone ring tones) have been measured in a variety of background sounds sampled from everyday listening situations, using a two-alternative forced-choice task. The different signals and backgrounds were interleaved, to create stimulus uncertainty, as would occur in everyday life. The data are used to relate the detectability index, d', to the calculated partial loudness. In this way, the model can be used to predict the detectability of any signal, based on its calculated partial loudness. [Work supported by MRC (UK) and by Nokia.]

9:30

5aPPb4. Sinusoidal spectral modulation masking period patterns. Aniket Saoji and David A. Eddins (Psychoacoustic Lab., Ctr. for Hearing and Deafness, Dept. of Commun. Disord. and Sci., Univ. at Buffalo, Buffalo, NY)

Knowledge of the internal auditory representation of spectral envelope features is critical to the understanding of auditory perception by listeners with normal and impaired hearing. Previous investigators have used forward and simultaneous masking to derive masking patterns that reveal the influence of lateral suppression and other factors on the internal representation of various spectral envelopes (e.g., narrowband noise and steady-state vowels). The present study followed a general approach, in which the internal representation of spectral shape was estimated on the basis of spectral masking patterns obtained with sinusoidal spectral modulation. Simultaneous- and forward-masked thresholds were measured for probe signals as a function of spectral modulation frequency (cycles/octave), phase, and depth (peak-valley contrast in dB) superimposed on a noise carrier (400–6400 Hz). For relatively low modulation frequencies (<1 cycle/octave), masking patterns closely mimicked features of the external spectrum. For very high modulation frequencies (>2 cycles/octave), masking patterns revealed reduced spectral contrast, consistent with the limited frequency selectivity of the auditory system. For intermediate modulation frequencies (1–2 cycles/octave), considerable spectral sharpening was evident. The three-dimensional space defined by spectral modulation frequency, depth, and internal contrast will be discussed in terms of predicting the internal representation of arbitrary complex spectral envelopes.

9:45

5aPPb5. The effect of modulation maskers on the detection of second-order amplitude modulation in the absence and presence of notched noise. Rosalie M. Uchanski (Dept. of Otolaryngol., Washington Univ. School of Medicine, St. Louis, MO 63110), Brian C. J. Moore, and Brian R. Glasberg (Univ. of Cambridge)

Second-order amplitude modulation (AM) can be detected despite the absence of a component at the second-order modulation frequency in the modulation spectrum. This may depend on the detection of a modulation component at the second-order frequency produced by nonlinearities in the auditory system, or on the detection of a ''beat'' at the output of the modulation filter centered at the first-order rate. To assess these possibilities, and to explore the role of off-frequency listening, first- and second-order AM detection thresholds were determined for six listeners with normal hearing, in the absence and presence of a notched noise centered at the carrier frequency of 4 kHz. First-order thresholds were measured for $f1 = 2$, 10, 16, and 50 Hz. Second-order thresholds were measured for $f1 = 16$ and $f2 = 2$ Hz, and for $f1 = 50$ and $f2 = 10$ Hz. Then, using a second-order modulation depth giving about 80%-correct detection, performance for detecting second-order AM was measured in the presence of a modulation masker centered at either the first- or second-order rate. In the absence of notched noise, both modulation maskers impaired performance, for both second-order rates. In the presence of notched noise, for most listeners there was little effect of either modulation masker. [Work supported by MRC.]

10:00

5aPPb6. Auditory discrimination of binary sequences. Stanley Sheft and William A. Yost (Parmly Hearing Inst., Loyola Univ. Chicago, 6525 N. Sheridan Rd., Chicago, IL 60626, ssheft@luc.edu)

Auditory sequence discrimination was investigated for stimuli generated with binary modulation coding schemes. Both amplitude- and frequency-shift keying (ASK and FSK) were used to convert 16-event binary patterns into a series of pulses. Discrimination was measured as a function of the number of binary events that differed between the standard and comparison sequences with temporal locations of the deviant pulses randomly selected on each trial. Across conditions, pulse duration ranged from 4 to 64 ms. In the ASK scheme, the carrier was wideband noise with off-key amplitudes scaled by a factor ranging from 0.0 to 0.5. Performance improved with increasing either pulse duration, amplitude shift, or the number of deviant pulses from one to eight. These trends are consistent with envelope processing by a modulation filterbank which discards phase information above low modulation rates. In the FSK scheme, binary patterns were coded by pulse-frequency differences centered at 1.25 kHz. As with ASK, FSK results showed performance improving with pulse duration and number of deviant pulses. FSK findings suggest that the low-pass results obtained with ASK envelope coding may reflect a general characteristic of sequence discrimination that is not restricted solely by auditory envelope processing. [Work supported by NIDCD.]

10:15

5aPPb7. Perceptually related analysis of time-frequency patterns via a hearing model (Sottek), a pattern-measurement algorithm (''relative approach'') and a window-deconvolution algorithm. Wade Bray (HEAD acoustics, Inc., 6964 Kensington Rd., Brighton, MI 48116)

The presence of patterns (time structure and/or tonal structure) in sound situations elicits a different analytic process by the human hearing system than when sound situations are without pattern. In the presence of pattern, two issues can interfere with the effectiveness of conventional measurement methods: the human hearing is not acting as an absolute magnitude measurer so conventional measurements may not, or only poorly, resolve the important criteria; and pattern information that is subjectively highly significant is frequently of low magnitude amid general high-level structure that is not a direct part of the subjective pattern-response problem yet carries significance. Several automotive sound quality and information technology time-data recordings involving various patterns, of different strengths, were analyzed using conventional, psychoacoustic, and specialized pattern-sensitive techniques. In general, the advanced perception-related methods proved superior in quantifying ''patterned'' sound situations according to subjective impression.

10:30

5aPPb8. Touch-induced scraping sounds and texture perception. M. Ercan Altinsoy (Inst. of Commun. Acoust., IGSN, Ruhr Univ. Bochum, 44780 Bochum, Germany, ercanaltinsoy@rub.de)

Sound is often the result of human–object interaction (touch, striking, scraping, etc.) and conveys to the listener required information about physical attributes of the interaction, and spatial properties of the sound event, e.g., location, geometry. People are highly skilled in using touch-produced sounds to identify texture properties. Loudness and pitch of the touch-induced scraping sounds are the important psychoacoustical determinants of the texture perception [S. J. Lederman, Perception **8**, 93–103 (1979)]. In this study, psychophysical experiments were conducted to in-

vestigate the relationship between auditory properties (loudness and pitch) and texture perception. In order to control the psychoacoustical parameters, the scraping sounds were synthesized in computer environment. The results of the experiments show that subjects were able to differentiate the roughness of textures on the basis of pitch and loudness differences. Decreasing pitch tends to result in increasing magnitude of perceived roughness.

10:45

5aPPb9. Revealing criteria which underlie judgments of auditory unpleasantness. Karin Zimmer, Wolfgang Ellermeier, and Christian Schmid (Dept. of Acoust., Aalborg, Univ., Fredrik Bajers Vej 7 B-5, DK-9440 Aalborg Denmark, kaz@acoustics.dk)

The usefulness of probabilistic choice models in attaining metric scales of sensation magnitude is demonstrated by scaling the perceived unpleasantness of environmental sounds. Seventy-four subjects made pairwise comparisons of the unpleasantness of 12 binaurally recorded sounds. The stimuli varied considerably in loudness, timbre, and temporal envelopes. A preference-tree choice model accounted well for the structure underlying the data, indicating that subjects changed criteria when evaluating different sound pairs, and that these criterion changes combined in a lawful way, making it possible to measure unpleasantness on a ratio scale across the entire set of sounds investigated. The sounds could be grouped according to their (nonacoustical) intrusiveness, and loudness. A subsequent multiple-regression analysis showed that in the subgroups identified, a combination of psychoacoustical sharpness and roughness explained the unpleasantness-scale values very well; $r^2_{\text{corr}} = 0.96$. It is thus illustrated that choice models provide an explicitly stated theory of the observers' decision behavior, with built-in checks of the consistency of judgments, and statistical tests to validate if the attempt at (ratio) scale construction succeeded. This is in contrast to techniques of ''direct'' scaling, e.g., magnitude estimation, in which the scale level has to be assumed *bona fide*. [Work supported by Centercontract on Sound Quality, Aalborg University.]

11:00

5aPPb10. Coherence and the Speech Intelligibility Index. James Kates (GN ReSound, Boulder Res. Group, 3215 Marine St., Rm. W161, Boulder, CO 80309, jkates@gnresound.dk) and Kathryn Arehart (Univ. of Colorado, Boulder, CO 80309)

Noise and distortion reduce the sound quality in hearing aids, but there is no established procedure for calculating sound quality in these devices. This presentation introduces a new intelligibility and sound-quality calculation procedure based on the Speech Intelligibility Index [ANSI S3.5-1997]. The SII involves measuring the signal-to-noise ratio (SNR) in separate frequency bands, modifying the estimated noise levels to include auditory masking, and computing a weighted sum across frequency of the modified SNR values. In the new procedure, the estimated signal and noise levels are replaced with estimates based on the coherence between the input and output signals of the system under test. Coherence is unaffected by linear transformations of the input signal, but is reduced by nonlinear effects such as additive noise and distortion; the SII calculation is therefore modified to include nonlinear distortion as well as additive noise. For additive noise, the coherence calculation gives SII scores identical to those computed using the standard procedure. Experiments with normal-hearing listeners using additive noise, peak-clipping distortion, and center-clipping distortion are then used to relate the computed coherence SII scores with the subjects' intelligibility and quality ratings. [Work supported by GN ReSound (JMK) and the Whitaker Foundation (KHA).]

FRIDAY MORNING, 28 MAY 2004 IMPERIAL BALLROOM B, 8:00 A.M. TO 12:00 NOON

Session 5aSC

Speech Communication: Poster Session IV

Lori L. Holt, Chair
Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

Contributed Papers

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and contributors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

5aSC1. The discrimination and the production of English vowels by bilingual Spanish/English speakers. Sandra Levey (Dept. of Speech-Lang.-Hearing Sci., Lehman College of the City Univ. of New York, 250 Bedford Park Blvd. W., Bronx, NY 10468)

The discrimination of English vowels in real and novel words by 40 bilingual Spanish/English participants was examined. Their discrimination was compared with that of 40 native monolingual English participants. Participants were 23–36 years of age (mean 25.3; median 25.0). Stimuli were presented within triads in an ABX paradigm. This categorial discrimination paradigm was selected to avoid labeling, allowing participants to indicate categories to which stimuli belonged. Bilingual participants' productions of vowels in real words used in the discrimination task were judged by two independent listeners. The goal was to determine the degree of correlation between discrimination and production. Vowels were studied as these segments present second language learners with more difficulty than consonants. Discrimination difficulty was significantly greater for bilingual participants than for native English participants for vowel contrasts and novel words. Significant errors also appeared in the bilingual participants' production of certain vowels. English vowels absent from Spanish presented the greatest difficulty, while vowels similar to those in Spanish presented the least difficulty. Earlier age of acquisition, absence of communication problems, and greater percentage of time devoted to communication in English contributed to greater accuracy in discrimination and production. [Work supported by PSC-CUNY.]

5aSC2. Proficient bilinguals require more information for vowel identification than monolinguals. Alexandra S. Lopez and Catherine L. Rogers (Dept. of Commun. Sci. & Disord., Univ. of South Florida, 4202 E. Fowler Ave., PCD1017, Tampa, FL 33620-8100)

Even proficient bilinguals have been shown to experience more difficulty understanding speech in noise than monolinguals. One potential explanation is that bilinguals require more information than monolinguals for phoneme identification. We tested this hypothesis by presenting gated, silent-center vowels to two groups of listeners: (1) monolingual American English speakers and (2) proficient Spanish–English bilinguals, who spoke unaccented or mildly accented English. To create the stimuli, two American English speakers were recorded as they read the following items: ''beeb, bibb, babe, bebb, babb,'' and ''bob.'' Duration-preserved silent-center versions of three tokens of each item were created by retaining varying amounts of the CV and VC transitions (10, 20, 30, or 40 ms) and attenuating the remainder of the vowel center to silence. Duration-neutral versions of silent-center tokens were created by lengthening or shortening the silent portion to match the tokens vowel duration to the average for all the tokens. Listeners identified the unedited (full vowel), duration-preserved, and duration neutral silent-center tokens in a six-alternative forced-choice task. The two groups of listeners identified the unedited tokens with similar accuracy. In the silent-center conditions, however, the bilinguals identified the stimuli less accurately than the monolinguals. [Work supported in part by NIH-NIDCD Grant No. 1R03DC005561-01A1.]

5aSC3. Discrimination of synthesized English vowels by American and Korean listeners. Byunggon Yang (English Dept., Dongeui Univ., 24 Kayadong, Pusanjingu, Pusan 614-714, Republic of Korea)

This study explored the discrimination of synthesized English vowel pairs by 27 American and Korean, male and female listeners. The average formant values of nine monophthongs produced by ten American English male speakers were employed to synthesize the vowels. Then, subjects were instructed explicitly to respond to AX discrimination tasks in which the standard vowel was followed by another one with the increment or decrement of the original formant values. The highest and lowest formant values of the same vowel quality were collected and compared to examine patterns of vowel discrimination. Results showed that the American and Korean groups discriminated the vowel pairs almost identically and their center formant frequency values of the high and low boundary fell almost exactly on those of the standards. In addition, the acceptable range of the same vowel quality was similar among the language and gender groups. The acceptable thresholds of each vowel formed an oval to maintain perceptual contrast from adjacent vowels. Pedagogical implications of those findings are discussed.

5aSC4. Acoustic cues in the perception of second language speech sounds. Anna A. Bogacka (School of English, Adam Mickiewicz Univ., al. Niepodleglosci 4, 61-874 Poznan, Poland)

The experiment examined to what acoustic cues Polish learners of English pay attention when distinguishing between English high vowels. Predictions concerned the influence of Polish vowel system (no duration differences and only one vowel in the high back vowel region), salience of duration cues and L1 orthography. Thirty-seven Polish subjects and a control group of English native speakers identified stimuli from heed-hid and who'd-hood continua varying in spectral and duration steps. Identification scores by spectral and duration steps, and $F1/F2$ plots of identifications, were given as well as fundamental frequency variation comments. English subjects strongly relied on spectral cues (typical categorical perception) and almost did not react to temporal cues. Polish subjects relied strongly on temporal cues for both continua, but showed a reversed pattern of identification of who'd-hood contrast. Their reliance on spectral cues was weak and had a reversed pattern for heed-hid contrast. The results were interpreted with reference to the speech learning model [Flege (1995)], perceptual assimilation model [Best (1995)] and ontogeny phylogeny model [Major (2001)].

5aSC5. Vowel normalization for accent: An investigation of perceptual plasticity in young adults. Bronwen G. Evans and Paul Iverson (Dept. of Phonet. and Linguist., Univ. College London, Wolfson House, 4, Stephenson Way, London NW1 2HE, UK, bron@phon.ucl.ac.uk)

Previous work has emphasized the role of early experience in the ability to accurately perceive and produce foreign or foreign-accented speech. This study examines how listeners at a much later stage in language development—early adulthood—adapt to a non-native accent within the same language. A longitudinal study investigated whether listeners who had had no previous experience of living in multidialectal environments adapted their speech perception and production when attending university. Participants were tested before beginning university and then again 3 months later. An acoustic analysis of production was carried out and perceptual tests were used to investigate changes in word intelligibility and vowel categorization. Preliminary results suggest that listeners are able to adjust their phonetic representations and that these patterns of adjustment are linked to the changes in production that speakers typically make due to sociolinguistic factors when living in multidialectal environments.

5aSC6. Exploring the intelligibilty of foreign-accented English vowels when "English" is ill-defined. Rikke Louise Bundgaard-Nielsen and Ocke-Schwen Bohn (English Dept., Aarhus Univ., Aarhus, Denmark)

Many studies of foreign-accented speech have been conducted in *second* language settings in which learners are assumed to be exposed to a relatively homogeneous non-native sound system. However, *foreign* language learners, who learn an additional language in a setting where this language is not the primary medium of communication, are frequently exposed to a range of varieties of the target language which may differ considerably with respect to their sound systems. The present study examined and compared the intelligibility of English monophthongs produced by two speaker groups: Native Danes who had learned English as a foreign language (with exposure to different native and non-native varieties) and native English speakers from Australia, the US, and the UK. Ten native Canadian-English listeners, who were familiar with native and non-native accents of English, identified the 11 monophthongs of English produced by the speaker groups in a /bVt/ context. As expected, the listeners' error patterns were specific for each speaker group. However, reduced intelligibility was observed for much the same vowels irrespective of speaker group. Our results suggest that one source of problems in learning the sounds of English is the heterogeneity of English vowel systems in addition to transfer from the native language.

5aSC7. ERP indices of speech processing in bilinguals. Karen Garrido, Miwako Hisagi, and Valerie Shafer (Speech and Hearing Sci., City Univ. of New York–Grad. Ctr., 365 Fifth Ave., New York, NY 10016)

The specific aim of this project was to examine perceptual differentiation of the vowel contrast [I-E] by early versus late bilingual Spanish–English speakers, compared to monolingual English speakers. Neither of these phonetic segments occur as phonemes in Spanish while they constitute phonemic contrast in English. For Spanish listeners, these vowels may be heard as variants of the Spanish phonemes [i] and [e], respectively. The formation of perceptual categories may be affected by exposure to two languages in which the phonemic status of phonetic segments differs. In the current study, electrophysiological measures (mismatched negativity—MMN) and behavioral measures (AX discrimination and identification) were employed to examine perception of a nine-step continuum resynthesized from natural tokens. Preliminary results indicate that early bilinguals show somewhat more categorical perception than monolinguals (i.e.,

sharper identification boundaries) of the nine-step continuum. MMN results, obtained during a task in which subjects were not attending to the auditory input, show clear MMNs for monolinguals for both cross- and within-category pairs. To date, results for early bilinguals show less clear indication of preattentive discrimination. The role of language experience and the relationship between the behavioral performance and the electrophysiological measures will be discussed.

5aSC8. Structural equation modeling applied to the prediction of word identification accuracy. Sumiko Takayanagi, Lynne E. Bernstein, and Edward T. Auer, Jr. (House Ear Inst., Los Angeles, CA 90057)

Five structural equation models (SEMs) were developed to examine explicitly the relative contribution of experiential (subjective frequency and word age of acquisition) and word similarity (lexical equivalence class size and phoneme equivalence class size) factors on word identification accuracy (WIA). WIA (percent word correct scores and word cost scores) was measured for 184 monosyllabic words under five presentation conditions: (1) visual-only speech; (2) high intelligibility vocoded auditory-only speech; (3) low intelligibility vocoded auditory-only speech; (4) high intelligibility vocoded audiovisual speech; and (5) low intelligibility vocoded audiovisual speech. The results showed that each factor can be treated as an isolated factor and can be measured explicitly. Furthermore, the relative strengths of their contributions varied as a function of intelligibility and the estimation power of the predictor variables. In addition, two models were developed to estimate audiovisual (conditions 4 and 5) WIA from experiential, visual- and auditory-similarity factors. Only the SEM of condition 4 fit the data, and the experiential and auditory factors contributed equally, but the visual factor did not contribute much to the WIA in this model. Advantages and limitations of SEM will be discussed.

5aSC9. Perceptual assimilation of French and German vowels by American English monolinguals: Acoustic similarity does not predict perceptual similarity. Winifred Strange, Erika Levy, and Robert Lehnholf, Jr. (Speech and Hearing Sci., City Univ. of New York—Grad. Ctr., 365 Fifth Ave., New York, NY 10016)

Previous research in our laboratory has demonstrated that the perceived similarity of vowels across languages is not always predictable from the closeness of their target formant values in $F1/F2/F3$ space. In this study, perceptual similarity was established using a task in which 11 American English (AE) monolinguals were presented multiple tokens of 9 French vowels and 14 North German vowels (in separate blocks) produced in citation-form /hVb(a)/ (bi)syllables by native speakers. They selected 1 of 11 AE vowel responses to which each non-native vowel token was most similar, and rated its goodness on a 9-point Likert scale. Of special interest was the perceptual assimilation of front rounded French [y, oe] and German [y, Y, o/, oe] vowels. Acoustically, all six French and German vowels are more similar to front unrounded AE vowels. However, all six vowels were perceived to be more similar to back rounded AE vowels (range across vowels = 55% to 100%), although relatively poor exemplars. There were differences across languages in how the same vowel was assimilated (e.g., French /y/ assimilated to front AE vowels 13%, German /y/, 0%; French [oe] 3%, German [oe] 45%). There were also large individual differences in listeners assimilation patterns. [Work supported by NIDCD.]

5aSC10. Segmental differences in the visual contribution to speech intelligibility. Kuniko Nielsen (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095-1543, kuniko@humnet.ucla.edu)

It is well known that the presence of visual cues increases the intelligibility of a speech signal (Sumby and Pollack, 1954). Although much is known about segmental differences in visual-only perception, little is known about the contribution of visual cues to auditory–visual perception for individual segments. The purpose of this study was to examine (1) whether segments differ in their visual contribution to speech intelligibility, and (2) whether the contribution of visual cues is always to increase speech intelligibility. One talker produced triples of real words containing 15 different English consonants. Forced-choice word-identification experiments were carried out with these recordings under auditory–visual (AV) and auditory-only (A) conditions with varying S/N ratios, and identification accuracy for the 15 consonants was compared for A versus AV conditions. As expected, there were significant differences in the visual contribution for the different consonants, with visual cues greatly improving speech intelligibility for most segments. Among them, labio-dentals and interdentals show the largest improvement. Although individual perceivers differed in their performance, the results also suggest that for some consonants, the presence of visual cues can reduce intelligibility. In particular, the intelligibility of [r] decreased significantly in the AV condition, being perceived as [w] in most cases.

5aSC11. The effects of age on identification of temporally altered visual-only speech signals. Brent Spehar, Nancy Tye-Murray (Dept. of Otolaryngol., Central Inst. for the Deaf, Washington Univ. School of Medicine, Box 8115, St. Louis, MO 63130), and Mitchell S. Sommers (Washington Univ., St. Louis, MO 63130, msommers@artsci.wustl.edu)

The purpose of the present study was to investigate the effects of age on the ability to identify temporally altered visual speech signals (i.e., visual-only presentations). Young (18–26) and older (over age 65) adults identified words placed in final position within a carrier phrase. Digital signal processing was used to compress (speed) and expand (slow) both the carrier phrase and the target words by 33%. Thus, the resultant stimuli were either 67% or 133% of the original duration. Speaking rate (expanded, unmodified, and compressed) was manipulated within subjects. Overall performance for both older and younger adults was significantly reduced for the speeded, compared with either the unaltered or the slowed condition. Furthermore, although older adults had significantly lower speechreading scores than younger adults, they did not exhibit disproportionate declines for the temporally altered conditions. These results stand in contrast to previous findings with auditory-only presentations in which older adults show significantly greater declines than younger adults as a consequence of increased speaking rate. Taken together, these findings argue against a generalized slowing of information processing in older adults and instead point to modality-specific changes in temporal processing abilities.

5aSC12. The effects of auditory–visual vowel and consonant training on speechreading performance. Carolyn Richie and Diane Kewley-Port (Dept. of Speech & Hearing Sci., Indiana Univ., 200 S. Jordan Ave., Bloomington, IN 47405, carodavi@indiana.edu)

Recent work examined the effects of a novel approach to speechreading training using vowels, for normal-hearing listeners tested in masking noise [C. Richie and D. Kewley-Port, J. Acoust. Soc. Am. **114**, 2337 (2003)]. That study showed significant improvements in sentence-level speechreading abilities for trained listeners compared to untrained listeners. The purpose of the present study was to determine the effects of combining vowel training with consonant training on speechreading abilities. Normal-hearing adults were tested in auditory–visual conditions in noise designed to simulate a mild-to-moderate sloping sensorineural hearing loss. One group of listeners received training on consonants in monosyllable context, and another group received training on both consonants and vowels in monosyllable context. A control group was tested but did not receive any training. All listeners performed speechreading pre- and post-tests, on words and sentences. Results are discussed in terms of differences between groups, dependent upon which type of training was administered; vowel training, consonant training, or vowel and consonant training combined. Comparison is made between these and other speechreading training methods. Finally, the potential benefit of these

vowel- and consonant-based speechreading training methods for rehabilitation of hearing-impaired listeners is discussed. [Work supported by NIHDCD02229.]

5aSC13. Visual contributions to talker attunement. Lawrence Brancazio (Dept. of Psych., Southern Connecticut State Univ., 501 Crescent St., New Haven, CT 06515 and Haskins Labs, New Haven, CT, brancazioL1@southernct.edu)

Listeners can attune to talker-specific speech patterns, as demonstrated by findings of greater accuracy in spoken-words-in-noise identification for words spoken by familiar voices [e.g., Nygaard and Pisoni, Percept. Psychophys. **60**, 355–376 (1998)]. In previous demonstrations of this effect, talker familiarity was produced by auditory speech exposure. The purpose of the present study was to examine the contribution of visible speech information to talker attunement. First, the auditory attunement effect was replicated: Participants were given extensive exposure to auditory words spoken by different talkers, and were subsequently tested on novel words in noise spoken by one of the familiar talkers and by a talker not previously heard. Identification of the words was more accurate for the familiar talker. This procedure was then extended to include video presentations of the talker's articulating face along with each word during the exposure phase. To control for familiarity with the faces, a separate condition included presentation of static images of the talkers' faces. Visual articulations were not presented during the identification-in-noise task. Thus, the experiment tests whether visual exposure to talker's articulatory patterns contributes to talker attunement, and, specifically, whether any such visual contributions to talker attunement generalize to auditory speech perception. [Work supported by NICHD.]

5aSC14. The influence of task on gaze during audiovisual speech perception. Julie Buchan, Martin Paré, Micheal Yurick, and Kevin Munhall (Queen's Univ., Kingston, ON, Canada, buchanj@psyc.queensu.ca)

In natural conversation, visual and auditory information about speech not only provide linguistic information but also provide information about the identity and the emotional state of the speaker. Thus, listeners must process a wide range of information in parallel to understand the full meaning in a message. In this series of studies, we examined how different types of visual information conveyed by a speaker's face are processed by measuring the gaze patterns exhibited by subjects watching audiovisual recordings of spoken sentences. In three experiments, subjects were asked to judge the emotion and the identity of the speaker, and to report the words that they heard under different auditory conditions. As in previous studies, eye and mouth regions dominated the distribution of the gaze fixations. It was hypothesized that the eyes would attract more fixations for more social judgment tasks, rather than tasks which rely more on verbal comprehension. Our results support this hypothesis. In addition, the location of gaze on the face did not influence the accuracy of the perception of speech in noise.

5aSC15. Cues to gender in children's speech. Suzanne Curtin and Scott Kiesling (Dept. of Linguist., Univ. of Pittsburgh, Pittsburgh, PA 15260, scurtin@pitt.edu)

Awareness of one's own gender emerges around 3 years and awareness that gender stays stable throughout life is evident by 4 years (Bee, 1998). This suggests that by 4 years of age, noticeable gender differences may emerge along a number of dimensions. The hypothesis tested here is that adults are able to identify the gender of 4-year-olds by voice quality alone. Sixteen four-year-olds were recorded saying the alphabet. Small portions of each recording were excised, and played to 40 adults. Adults were asked to identify the gender of the speaker. Subjects were able to correctly identify the gender of the child more often than chance. However, in the cases where a child's gender was incorrectly identified, pitch did not play a significant role. Rather, it appears that formant structure is the best predictor (although not perfect), of how a child's gender will be judged by voice (Perry, 2001). These results have important implications for our understanding of the linguistic cues that listeners use to identify the gender of speakers; they must be relying on phonetic cues that are much more subtle than gross pitch, lexical, phonological, or syntactic differences which are the usual provisions of language and gender research.

5aSC16. Role of a carrier sentence in native speakers' identification of Japanese vowel length. Yukari Hirata (Dept. of EALL, 13 Oak Dr., Hamilton, NY 13346, yhirata@mail.colgate.edu) and Stephen G. Lambacher (Univ. of Aizu, Aizu-Wakamatsu, Fukushima 965-8580, Japan)

This study examined the role of word-external contexts in native speakers' identification of phonemic vowel length in Japanese disyllabic nonwords. Three groups of native Japanese listeners ($n = 20$, 20, and 19) were assigned to three conditions differing in word-external contexts. In intact condition, three types of target disyllables /mVmV/, /mV:mV/, and /mVmV:/ were spoken in a carrier sentence at slow and fast speaking rates. In the other conditions, the target disyllables of the intact condition were excised from the original carrier sentence (excised), and embedded in the carrier sentence of the other rate (mismatch). The accuracy for identifying the length of the two vowels was significantly higher for the intact (98.1%) than the excised (69.5%) and the mismatch condition (52.4%). This indicates that the presence of the carrier sentence with an appropriate rate was essential for accurate identification of vowels, consistent with Johnson and Strange [J. Acoust. Soc. Am. **72**, 1761–1770 (1982)]. Further analysis indicated that the perceptual errors made in the excised and the mismatch conditions were closely associated with the absolute duration overlap between the short vowels spoken at the slow rate and the long vowels spoken at the fast rate.

5aSC17. On the just-noticeable difference for tempo in speech. Hugo Quené (Utrecht Inst. of Linguist., Utrecht Univ., The Netherlands, hugo.quene@let.uu.nl)

The rate or tempo of speech modulates listeners' phonetic expectations, e.g., about VOT or about segment durations. In addition, variations in tempo contribute directly to speech communication, e.g., by expressing the communicative importance of speech fragments. It is not clear, however, how large tempo differences in speech have to be in order to be noticeable and to contribute to speech perception. This paper attempts to determine the jnd for tempo in speech, by means of detection of gradual changes in tempo. Participants listened to 10-s spoken texts, in which overall tempo was linearly accelerated (to 80%) or decelerated (to 120%) over a 5-s interval. They were instructed to press a button as soon as they detected a tempo change in the text; jnd's were computed from their response times. Preliminary results suggest that the average jnd's for tempo in speech are about 14% (acceleration) and 15% (deceleration). These values are far larger than those reported for tempo in music (about 3% to 6%), which suggests that only large changes in tempo are relevant for speech communication.

5aSC18. Individual differences in speech and nonspeech perception of frequency and duration. Matthew J. Makashay[a] (Dept. of Linguist., Ohio State Univ., 222 Oxley Hall, 1712 Neil Ave., Columbus, OH 43210, matthew.makashay@na.amedd.army.mil)

This study investigates whether there are systematic individual differences in the perceptual weighting of frequency and durational speech cues for vowels and fricatives (and their nonspeech analogs) among a dialectally homogeneous group of speakers. Listeners performed AX discrimination for four separate types of stimuli: sine wave vowels, narrow-band fricatives, synthetic vowels, and synthetic fricatives. Duration and $F1$ frequency were manipulated for the vowels in *heed* and *hid*, and duration and

frequency of the fricative centroid in the *F*5 region were manipulated for the fricatives in *bath* and *bass*. Dialect production and perception tasks were included to ensure that subjects were not from dissimilar dialects. Multidimensional scaling results indicated that there are subgroups within a dialect that attend to frequency and duration differently, and that not all listeners use these cues consistently across dissimilar phones. If subgroups can have different perceptions of speech despite similar productions, this questions the requirements for classifying dialect continua. Furthermore, the ratios of these subgroups changing over time can explain some language mergers and shifts. [a] Currently at Army Audiol. & Speech Ctr., Walter Reed Army Medical Ctr., Washington, D.C.

5aSC19. Acoustic–phonetic convergence among interacting talkers. Jennifer S. Pardo (Psych. Dept., Columbia Univ., 1190 Amsterdam Ave. Mail Code 5501, New York, NY 10027, jsp2003@columbia.edu)

Among other sources, acoustic–phonetic variability derives from both personal attributes of a talker and the social structure of a conversational setting. This research begins to detail the effects of such sources by examining the influence of conversational interaction on speech production. A set of talkers provided samples of speech before, during, and after participating in a paired conversational task. Using an A×B paradigm with a separate set of listeners, initial research found that paired talkers became more similar in phonetic repertoire, that this convergence persisted over a short delay, and that the degree of convergence was influenced by both the sex and task role of a talker. This paper presents detailed acoustic and talker scaling analyses from this conversational corpus in an attempt to chart the effects of personal and situational factors on speech variability. [Research supported in part by NIMH Grant No. 5F32MH64995.]

5aSC20. Effects of speaking rate on the acceptability of change in segmental duration within a phrase. Makiko Muto,[a] Hiroaki Kato (ATR Human Information Sci. Labs., Kyoto 619-0288, Japan, makiko.muto@ruri.waseda.jp), Minoru Tsuzaki (ATR Spoken Lang. Translation Res. Labs., Kyoto 619-0288, Japan), and Yoshinori Sagisaka (Waseda Univ., Tokyo 169-0051, Japan)

To contribute to the naturalness criteria of speech synthesis, acceptability of changes in segment duration has been investigated. Previous studies showed context dependency of the acceptability evaluation such as *intraphrase positional effect*, where listeners were more sensitive to the phrase-initial segment duration than the phrase-final one. Such contextual effects were independent of the original durations of the segments tested [Kato *et al.*, J. Acoust. Soc. Am. **104**, 540–549 (1998)]. However, past studies used only normal-speed speech and temporal variation was limited. The current study, therefore, examined the contextual effect with a wide variety of speaking rates. The materials were three-mora phrases with either rising or falling accent that were spoken at three rates (fast, normal, and slow) with or without a carrier sentence. The duration of each vowel was either lengthened or shortened (10–50 ms) and listeners evaluated the acceptability of these changes. The results showed a clear speaking-rate effect in parallel with the intraphrase positional effect: the acceptability declined more rapidly as the speaking rate became faster. These results, along with those of Kato *et al.*, suggest that acceptability is evaluated based on the speaking rate rather than on the original duration itself. [Work supported by TAO, Japan.] [a] Currently at GITI, Waseda University.

5aSC21. Perceptual learning of talker-specific vowel space. Jennifer S. Queen (Dept. of Psych., Rollins College, Winter Park, FL 32789, jqueen@rollins.edu) and Lynne C. Nygaard (Emory Univ., Atlanta, GA 30322)

Previous research suggests that as listeners become familiar with a speaker's vocal style, they are better able to understand that speaker. This study investigated one possible mechanism by which this talker familiarity benefit arises. Listeners' vowel spaces were measured using a perceptual discrimination test both before and after they were trained to identify a group of speakers by name. Listeners identified either the same speakers whose vowels they discriminated or a different group of speakers. Differences in the learnability and the intelligibility of the two speaking groups were observed. The speaker group that was harder to identify also had vowels that were harder to discriminate. Changes in the listeners' vowel spaces were determined by examining multidimensional scaling solutions of their responses during the discrimination tests. All listeners became better at discriminating vowels. However, only listeners who heard different speakers during identification training and vowel discrimination exhibited a shift in their vowel spaces after training. This suggests that encountering new voices in an unrelated, nonlinguistic task acts to alter the perceptual context and may affect the structure of linguistic representation. Together, these results suggest a link between linguistic and nonlinguistic information in representations for spoken language.

5aSC22. Talker-specific auditory imagery during reading. Lynne C. Nygaard, Jessica Duke, Kathleen Kawar (Dept. of Psych., Emory Univ., Atlanta, GA 30322, lnygaar@emory.edu), and Jennifer S. Queen (Rollins College, Winter Park, FL 32789)

The present experiment was designed to determine if auditory imagery during reading includes talker-specific characteristics such as speaking rate. Following Kosslyn and Matt (1977), participants were familiarized with two talkers during a brief prerecorded conversation. One talker spoke at a fast speaking rate and one spoke at a slow speaking rate. During familiarization, participants were taught to identify each talker by name. At test, participants were asked to read two passages and told that either the slow or fast talker wrote each passage. In one condition, participants were asked to read each passage aloud, and in a second condition, they were asked to read each passage silently. Participants pressed a key when they had completed reading the passage, and reading times were collected. Reading times were significantly slower when participants thought they were reading a passage written by the slow talker than when reading a passage written by the fast talker. However, the effects of speaking rate were only present in the reading-aloud condition. Additional experiments were conducted to investigate the role of attention to talker's voice during familiarization. These results suggest that readers may engage in auditory imagery while reading that preserves perceptual details of an author's voice.

5aSC23. Duration as a cue in the distinction between glides and vowels. A. E. Coren (Dept. of Psych., Univ. of Texas, 1 University Station A8000, Austin, TX 78712-0187, aecoren@mail.utexas.edu) and W. J. Warren (Univ. of Texas, Austin, TX 78712-0198)

This experiment examines the role of length in the perception of glides and vowels. Tokens of the word "rodeo," spoken at two speech rates, were derived from an original by taking one cycle out of the waveform of the high vowel. Twenty listeners were presented with these derived tokens. It was found that when the length of the first vowel was approximately half that of the second, then 50 percent of the listeners judged the segment a vowel and 50 percent judged the segment a glide. This study agrees with earlier work done in this same manner. It is concluded that in running speech, the relative length of the first vowel is crucial in the vowel versus glide perceptual distinction.

5aSC24. Perceptual effects of preceding non-speech-rate information on temporal properties of speech categories. Travis Wade and Lori Holt (Dept. of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, twade@andrew.cmu.edu)

The purpose of the study was to determine whether a general mechanism such as durational contrast can explain the apparent rate-dependent nature of speech perception. Four experiments were performed to test

effects of pure tone presentation rate on perception of following speech continua involving duration-varying formant transitions with which the tones shared critical temporal and spectral properties. Results showed small but consistent shifts in the stop-continuant boundary distinguishing /ba/ and /wa/ syllables based on the rate of precursor tones similar in duration and frequency to syllable-initial $F1$ and $F2$ transitions, across differences in amplitude of tones and despite variability in their duration. Additionally, the shift was shown to involve the entire graded structure of the [w] category and was not limited to an ambiguous boundary region, affecting goodness judgments on both sides of an estimated best exemplar range. These results are problematic for accounts of rate-dependent processing that explicitly reference speech categories or articulatory events and are consistent with a contrast account.

5aSC25. Medial surface dynamics of the vocal folds in an *in vivo* canine model. Michael Doellinger, Gerald S. Berke, Dinesh K. Chhetri, and David A. Berry (Div. of Head and Neck Surgery, UCLA School of Medicine, 31-24 Rehab., 1000 Veteran Ave., Los Angeles, CA 90095-1794)

Quantitative measurement of the medial surface dynamics of the vocal folds is important for understanding how sound is generated in the larynx. However, such data are hard to gather because of the inaccessibility of the vocal folds. Recent studies have applied hemi-larynx methodology to excised human larynges, to visualize these dynamics. The present study extends this methodology to obtain similar quantitative measurements using an *in vivo* canine hemi-larynx setup, with varying levels of stimulation to the recurrent laryngeal nerve. Use of an *in vivo* model allows us to examine effects of intrinsic muscle contraction on the medial surface of the vocal folds, to provide greater insight into mechanisms of vocal control. Data were collected using digital high-speed imaging with a sampling frequency of up to 4000 Hz, and a spatial resolution of up to 1024 $\times$1024 pixels. Three-dimensional motion will be extracted, computed, visualized, and contrasted as a function of the level of stimulation to the recurrent laryngeal nerve. Results will also be compared to patterns of vibration in excised larynges. Finally, commonly applied quantitative analyses will be performed to investigate the underlying modes of vibration. [Work supported by NIH/NIDCD.]

5aSC26. Perceptual relevance of source spectral slope measures. Jody Kreiman and Bruce R. Gerratt (Div. of Head/Neck Surgery, UCLA School of Medicine, 31-24 Rehab Ctr., Los Angeles, CA 90095-1794, jkreiman@ucla.edu)

Researchers broadly agree that the spectral slope of the voice source is an important concomitant of voice quality. Many measures of source spectral slope have been proposed, including the relative amplitudes of the lowest few harmonics, the ratio of energy in low- versus high-frequency bands, and the average deviation from an ideal slope. It is unclear which (if any) of these measures best reflects the differences in vocal quality that result from the underlying acoustic variability. To examine this issue, a large corpus of voice samples was inverse filtered, and spectra were calculated for resulting source pulses. Different measures of spectral slope were calculated for each voice, and correlations among measures were examined. Finally, several series of synthetic stimuli were created in which only the source spectral slope varied in steps. Listeners judged the similarity of stimuli within each series. Similarity responses were evaluated with multidimensional scaling, and the resulting perceptual spaces were interpreted in terms of the different measures of source spectral slope. Measures that are highly correlated with the perceptual spaces reflect perceptually important aspects of the source signal. [Research supported by NIDCD.]

5aSC27. Significance of analysis window size in maximum flow declination rate. Linda M. Carroll (Dept. of Otolaryngol., Mount Sinai School of Medicine, 5 E. 98th St., 1st Fl., Box 1653, Grabscheid Voice Ctr., New York, NY 10029, linda.carroll@mssm.edu)

Most acoustic stability measures and aerodynamic measures have recommended data analysis methods to ensure accurate comparison across research studies. Although standard aerodynamic analysis methods exist for subglottal pressure and mean transglottal flow, such standards do not exist for maximum flow declination rate (MFDR). As such, it becomes difficult to compare results across studies with the wide range of analysis windows. Because MFDR is a strong indicator of laryngeal function and increasingly is being reported in aerodynamic studies, it is necessary to determine whether a significant difference exists between analysis windows in data management. This study finds significant differences in MFDR with comparison of four different data extraction methods on the same data set. The contrasting data extractions compare differences between method A (MFDR from entire 1000-ms segment, excluding onset/offset), method B (MFDR from middle 100 ms with center at midportion of entire segment), method C (MFDR from middle 100 ms with center at midportion of maximum MFDR value from entire segment), and method D (MFDR from 20 consecutive cycles with center at midportion of maximum MFDR value from entire segment) for a sustained 7-syllable /pa/ repetition at two fundamental frequencies.

5aSC28. Fundamental frequency perturbation indicates perceived health and age in male and female speakers. David R. Feinberg (School of Psych., Univ. of St. Andrews, St. Andrews, Fife KY169JU, Scotland, UK)

There is strong support for the idea that healthy vocal chords are able to produce fundamental frequencies ($F0$) with minimal perturbation. Measures of $F0$ perturbation have been shown to discriminate pathological versus healthy populations. In addition to measuring vocal chord health, $F0$ perturbation is a correlate of real and perceived age. Here, the role of jitter (periodic variation in $F0$) and shimmer (periodic variation in amplitude of $F0$) in perceived health and age in a young adult (males aged 18–33, females aged 18–26), nondysphonic population was investigated. Voices were assessed for health and age by peer aged, opposite-sex raters. Jitter and shimmer were measured with Praat software (www.praat.org) using various algorithms (jitter: DDP, local, local absolute, PPQ5, and RAP; shimmer: DDA, local, local absolute, APQ3, APQ5, APQ11) to reduce measurement error, and to ascertain the robustness of the findings. Male and female voices were analyzed separately. In both sexes, ratings of health and age were significantly correlated. Measures of jitter and shimmer correlated negatively with perceived health, and positively with perceived age. Further analysis revealed that these effects were independent in male voices. Implications of this finding are that attributions of vocal health and age may reflect actual underlying condition.

5aSC29. Spectral anlaysis of glottal flow models. Matthew E. Lee (Ctr. for Signal and Image Processing, Georgia Inst. of Technol., mattlee@ece.gatech.edu) and Mark J. T. Smith (Purdue Univ.)

An analysis of the effects of time-domain characteristics of glottal waveforms in the spectral domain is presented. Studies have shown that the relative difference in amplitude of the first two harmonics of the inverse-filtered voice waveform ($H1^*-H2^*$) can be used to predict the value of the open quotient in certain situations [J. Sundberg, M. Andersson, and C. Hultqvist, J. Acoust. Soc. Am. **105**, 1965–1971 (1999)]. However, recent experiments have shown that glottal waveform asymmetry is an additional factor that will affect the relationship between the harmonics and open quotient. This work is aimed at establishing a method in which both the open quotient and a glottal asymmetry coefficient can be reliably determined based on the characteristics of the first two harmonics. By deriving analytic formulas for the spectra of two glottal flow models (*LF*, $R++$), it is shown that a measure of the relative phase difference of the

first two harmonics as well as the amplitude difference can be used in order to determine these perceptually important time-domain parameters of the glottal waveform. A theoretical analysis as well as experimental results are presented.

5aSC30. An improved correction formula for the estimation of voice source harmonic magnitudes. Markus Iseli (Dept. of Elec. Eng., UCLA, 405 Hilgard Ave., Los Angeles, CA 90095, iseli@ee.ucla.edu) and Abeer Alwan (UCLA, Los Angeles, CA 90095)

Information about voice and talker characteristics, or voice quality, can be gained from glottal flow estimates. Many voice quality parameters, such as the open quotient (OQ), depend on an accurate estimate of the glottal flow (voice source) spectrum. It is known that OQ, for example, is correlated with the magnitude difference of the first two harmonics (H_1-H_2) of the voice source spectrum [Holmberg *et al.*, "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice," J. Speech Hear. Res. **38**, 1212–1223 (1995)]. To obtain an accurate estimate of the voice source spectral harmonics, the influence of vocal–tract resonances needs to be removed. In Hanson, Ph.D. dissertation, Harvard University (1995), a correction formula which removes the effect of F_1 on H_1 and H_2 was presented. In this work, an improved correction formula for the estimation of source harmonics′ magnitudes is presented. The new correction formula accounts for the bandwidths of vocal–tract resonances, and most importantly, is not limited to the analysis of nonhigh vowels. H_1-H_2 estimates, using the proposed technique with synthesized vowels /a/, /i/, and /u/ generated with the LF and the KLGLOTT88 models, are very accurate. [Work supported in part by the NSF.]

5aSC31. A comparison of harmonic production in trained and untrained singers. Jonathan H Hildebrand and Hilary J. Caso (Dept. of Linguist., Univ. of North Carolina, Chapel Hill, CB#3155, Chapel Hill, NC 27599, hildebr1@email.unc.edu)

This study was designed to investigate the differences in the tone qualities of trained and untrained singers of both genders by measuring the ratio of energy found in the fundamental frequency to that found in the harmonics of the two groups. Forty trained and untrained male and female subjects (ten in each group) sang the vowels [a e i o u] in the words "me cake father obey too." Each subject sang through the list of words in its entirety one time. They were also asked to produce each word at a constant pitch of their choosing. The intensity (dB) was then measured at the vowel midpoint in the intact vowel, and then again with the fundamental filtered out. These two measurements allowed for a comparison to be made by subtracting the results. This process produced results showing that the fundamental frequency was responsible for the majority of the trained singers', vocal productions, where harmonics were responsible for the majority of the untrained singers', vocal productions. The data also showed that these results were more prominent for the production of the high vowels in both genders, and for women overall.

5aSC32. Lexical frequency and voice assimilation in complex words in Dutch. Mirjam Ernestus, Mybeth Lahey, Femke Verhees, and Harald Baayen (Max-Planck Inst. for Psycholinguist. and Nijmegen Univ., Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands, mirjam.ernestus@mpi.nl)

Words with higher token frequencies tend to have more reduced acoustic realizations than lower frequency words (e.g., Hay, 2000; Bybee, 2001; Jurafsky *et al.*, 2001). This study documents frequency effects for regressive voice assimilation (obstruents are voiced before voiced plosives) in Dutch morphologically complex words in the subcorpus of read-aloud novels in the corpus of spoken Dutch (Oostdijk *et al.*, 2002). As expected, the initial obstruent of the cluster tends to be absent more often as lexical frequency increases. More importantly, as frequency increases, the duration of vocal-fold vibration in the cluster decreases, and the duration of the bursts in the cluster increases, after partialing out cluster duration. This suggests that there is less voicing for higher-frequency words. In fact, phonetic transcriptions show regressive voice assimilation for only half of the words and progressive voice assimilation for one third. Interestingly, the progressive voice assimilation observed for higher-frequency complex words renders these complex words more similar to monomorphemic words: Dutch monomorphemic words typically contain voiceless obstruent clusters (Zonneveld, 1983). Such high-frequency complex words may therefore be less easily parsed into their constituent morphemes (cf. Hay, 2000), favoring whole word lexical access (Bertram *et al.*, 2000).

5aSC33. An acoustic and electroglottographic study of V-glottal stop-V in two indigenous American languages. Christina M. Esposito and Rebecca Scarborough (Dept. of Linguist., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095, esposito@humnet.ucla.edu)

Both Pima, a Uto-Aztecan language spoken in Arizona, and Santa Ana del Valle Zapotec (SADVZ), an Otomanguean language spoken in Oaxaca, Mexico, have sequences of two vowels separated by an intervening glottal stop. In both languages, this V?V sequence becomes reduced in certain occurrences, with the perceptual effect of the loss of /?/ in Pima and the loss of $V2$ in SADVZ. The purpose of this study is to provide an acoustic and electroglottographic (EGG) description of these sequences in both their full and reduced forms, prompted by varying speech rate. Two acoustic measures of phonation type ($H1-H2$, $H1-A3$) and two EGG measures (OQ and peak closing velocity) were made at vowel midpoints and adjacent to /?/. For Pima, an issue of interest is what properties of the /V?V/ sequences (when $V1=V2$) allow them to be distinguished from phonemic long vowels in the reduced forms where /?/ is lost. It is hypothesized that /?/ will be preserved as vowel glottalization. For SADVZ, an important question is why the vowels sound creaky despite a lack of spectral evidence for creak. It is hoped that more direct EGG measures will show the perceived phonation.

5aSC34. Acoustic and aerodynamic characteristics of ejectives in Amharic. Didier Demolin (Universidade de Sao Paulo/CNPq, 403 av. Prof. Gualberto, 01060-970 Sao Paulo, Brazil)

This paper invetsigates the main phonetic characteristics that distinguishes ejectives from pulmonic sounds in Amharic. In this language, there are five ejectives that can be phonemically singleton or geminate. Duration measurements have been made in intervocalic position for pulmonic stops and for each type of ejective, taking into account the overall duration and VOT. Results show that ejective stops have a higher amplitude burst than pulmonic stops. The duration of the noise is shorter for ejective fricatives compared to pulmonic fricatives. At the end of ejective fricatives, there is a 30-ms glottal lag that is not present in pulmonic fricatives. Geminate ejectives are realized by delaying the elevation of the larynx. This can be observed on the spectrographic data by an increase of the noise at the end of the geminate ejectives. Aerodynamic data have been collected in synchronization with the acoustic recordings. The main observations are that pharyngeal pressures values are much higher than what is usually assumed (up to 40 CmH2O for velars) and that the delayed command in the elevation of the larynx of geminate ejectives is shown by two phases in the rise of pharyngeal pressure.

5aSC35. Aerodynamic characteristics of French consonants. Didier Demolin (Universidade de Sao Paulo, 403 av. Prof. Gualberto, 01060-970 Sao Paulo, Brazil), Sergio Hassid (Hopital Ersame, Universite libre de Bruxelles), and Alain Soquet (Universite Libre de Bruxelles, 1050 Brussels, Belgium)

This paper reports some aerodynamic measurements made on French consonants with a group of ten speakers. Speakers were recorded while saying nonsense words in phrases such as papa, dis papa encore. The

nonsense words in the study combined each of the French consonants with three vowels /i, a, u/ to from two syllables words with the first syllable being the same as the second. In addition to the audio signal, recordings were made of the oral airflow, the pressure of the air in the pharynx above the vocal folds and the pressure of the air in the trachea just below the vocal folds. The pharyngeal pressure was recorded via a catheter (i.d. 5 mm) passed through the nose so that its open end could be seen in the pharynx below the uvula. The subglottal pressure was recorded via a tracheal puncture between the first and the second rings of the trachea or between the cricoid cartilage and the first tracheal ring. Results compare subglottal presssure, pharyngeal pressure, and airflow values. Comparisons are made between values obtained with male and female subjects and various types of consonants (voiced versus voiceless at the same place of articulation, stops, fricatives, and nasals).

5aSC36. Viscoelastic properties of the false vocal fold. Roger W. Chan (Otolaryngol. and Grad. Prog. in Biomed. Eng., Univ. of Texas Southwestern Medical Ctr., Dallas, TX 75390-9035, roger.chan@utsouthwestern.edu)

The biomechanical properties of vocal fold tissues have been the focus of many previous studies, as vocal fold viscoelasticity critically dictates the acoustics and biomechanics of phonation. However, not much is known about the viscoelastic response of the ventricular fold or false vocal fold. It has been shown both clinically and in computer simulations that the false vocal fold may contribute significantly to the aerodynamics and sound generation processes of human voice production, with or without flow-induced oscillation of the false fold. To better understand the potential role of the false fold in phonation, this paper reports some preliminary measurements on the linear and nonlinear viscoelastic behavior of false vocal fold tissues. Linear viscoelastic shear properties of human false fold tissue samples were measured by a high-frequency controlled-strain rheometer as a function of frequency, and passive uniaxial tensile stress-strain response of the tissue samples was measured by a muscle lever system as a function of strain and loading rate. Elastic moduli (Young's modulus and shear modulus) of the false fold tissues were calculated from the measured data. [Work supported by NIH.]

5aSC37. Determining the relationship between respiratory variables and fundamental frequency declination: A mathematical modeling of the data. Carole E. Gelfer (Dept. of Commun. Disord., William Paterson Univ., Wayne, NJ 07470) and Jay Jorgenson (City College of New York, New York, NY 10031)

Previous research [C. E. Gelfer, unpublished doctoral dissertation, City University of New York Graduate Center (1987)] demonstrated stability in fundamental frequency ($F0$) declination across utterance of various lengths that appeared to be derived from a stable underlying subglottal pressure (Ps). Moreover, this stability was maintained even when airflow requirements were varied in response to the varying phonetic composition of utterances. However, although a preliminary mathematical analysis suggested that both the Ps and $F0$ could be characterized as second order linear systems, no formal testing of these curves was performed. In this study we compare these data employing the following statistical procedure. First, we hypothesize that the data from both sets can be modeled by a curve of the form $y=axbecx$, with unknown parameters a, b, and c which are computed. We then consider a multivariate confidence interval for the vector (a,b,c) of fitted parameters, as well as the difference of the vectors for two such models. The comparison is made through distributional estimates obtained via Monte Carlo simulation.

5aSC38. Flow characteristics in a scaled-up glottis model. Michael Barry, Timothy Wei (Dept. of Mech. and Aerosp. Eng., Rutgers Univ., Piscataway, NJ 08854), and Michael Krane (Rutgers Univ., Piscataway, NJ 08854)

Measurements of the velocity field of the flow through a scaled-up model of the vocal folds are presented. Scaling up the model ten times life size and using water as a working fluid enables temporally and spatially resolved measurements of the velocity using digital particle image velocimetry, making use of dynamic similarity. The Reynolds number (based on maximum glottal width and maximum jet speed) is 3000, and the Strouhal number (based on glottis length, maximum jet speed, and opening/closing time) ranges from 0.01 to 0.1. Prescribing the motion of the model vocal folds allows phase averaging of the flowfield. Measurements presented focus on the time intervals just after vocal fold opening and just before closure. [Work funded by NIDCD of NIH.]

5aSC39. Acoustic correlates of Japanese expressions associated with voice quality of male adults. Hiroshi Kido (Dept. of Commun. Eng., Tohoku Inst. of Technol., Sendai, Japan) and Hideki Kasuya (Utsunomiya Univ., Utsunomiya 321-8585, Japan)

Japanese expressions associated with the voice quality of male adults were extracted by a series of questionnaire surveys and statistical multivariate analysis. One hundred and thirty-seven Japanese expressions were collected through the first questionnaire and careful investigations of well-established Japanese dictionaries and articles. From the second questionnaire about familiarity with each of the expressions and synonymity that were addressed to 249 subjects, 25 expressions were extracted. The third questionnaire was about an evaluation of their own voice quality. By applying a statistical clustering method and a correlation analysis to the results of the questionnaires, eight bipolar expressions and one unipolar expression were obtained. They constituted high-pitched/low-pitched, masculine/feminine, hoarse/clear, calm/excited, powerful/weak, youthful/elderly, thick/thin, tense/lax, and nasal, respectively. Acoustic correlates of each of the eight bipolar expressions were extracted by means of perceptual evaluation experiments that were made with sentence utterances of 36 males and by a statistical decision tree method. They included an average of the fundamental frequency ($F0$) of the utterance, speaking rate, spectral tilt, formant frequency parameter, standard deviation of $F0$ values, and glottal noise, when SPL of each of the stimuli was maintained identical in the perceptual experiments.

5aSC40. Intrinsic fundamental frequency effects in hearing impaired speakers. Bryan Gick (Dept. of Linguist., Univ. of BC, 270-1866 Main Mall, Vancouver, BC V6T 1Z1, Canada), Barbara Bernhardt, and Penelope Bacsfalvi (Univ. of BC, Vancouver, BC)

The source of the well-known intrinsic fundamental frequency (IF0) effect of vowel height has been controversial for decades. Previous work has found the average IF0 effect cross-linguistically to be 15.3 Hz [Whalen and Levitt, J. Phonetics (1995)]. The present study investigates IF0 for four hearing-impaired speakers. Based on previous observations that profoundly hearing impaired speakers vary voice pitch less than normal hearing speakers [Osberger and McGarr, Speech Lang. (1982)], our participants were expected to show a reduced IF0 effect. However, results show an average IF0 effect of 22 Hz, with a markedly wide range across speakers, from -4 to 48 Hz, with three of the four participants showing an above average-sized effect. Further, results of measures taken following speech articulation intervention using visual feedback [Bernhardt *et al.*, Clin. Linguist Phonet. (2003)] show a decrease in IF0 for the speakers with an over-sized effect, and an increase in the speaker with an under-sized effect, despite that neither IF0 nor pitch in general were included in treatment. Results of this study support a lingual-articulatory origin for IF0, as well as suggesting that normal-hearing speakers may use auditory feedback to mediate what would otherwise be a larger effect. [Research supported by NSERC.]

5aSC41. The battle for *F*0: Glottalization versus stress in Statimcets. Sonya Bird and Marion Caldecott (Univ. of BC, 1866 Main Mall, Buchanan E270, Vancouver, BC V6T 1Z1, Canada)

The conflicts that arise in speech production give us insight into the interaction between faithfulness to mental representations and articulatory and acoustic limitations on speech. This paper presents an example of a perceptually based conflict [Kochetov, LabPhon8 (to appear)] in Statimcets, a Salish language spoken in the Interior of British Columbia. In a pilot study on glottalized resonants in Statimcets, Bird [ICSNL 38 (2003)] found that the primary cue to glottalization was creaky voicing, and that glottalization was reduced in stressed syllables. This paper expands on the previous study, incorporating new acoustic data on stress, and a more detailed acoustic analysis of glottalization. It is shown that one of the cues to stress, raised fundamental frequency ($F0$), conflicts with the lowered $F0$ associated with creaky voicing. In Statimcets, this conflict is resolved by preserving the perceptual salience of stress cues at the expense of cues to glottalization. This results in loss of glottalization in stressed syllables. Acoustic data and analyses will be presented. [Research supported by SSHRC.]

FRIDAY MORNING, 28 MAY 2004 CONFERENCE ROOM L, 8:55 TO 11:30 A.M.

Session 5aSP

Signal Processing in Acoustics: Selected Topics

Juliette Ioup, Chair
Physics Department, University of New Orleans, New Orleans, Louisiana 70148

Chair's Introduction—8:55

Contributed Papers

9:00

5aSP1. A multifilter approach to acoustic echo cancellation. John Usher, Wieslaw Woszczyk, and Jeremy Cooperstock (Ctr. for Interdisciplinary Res. in Music Media and Technol., McGill Univ., Montreal, QC, Canada, jusher@po-box.mcgill.ca)

Hands-free teleconferencing is increasingly frequent today. An important design consideration for any such communication tool that uses high-quality audio is the return echo caused by the acoustic coupling between the loudspeakers and microphones at each end of the conference. An echo-suppression filter (ESF) reduces the level of this return echo, increasing speech intelligibility. A new ESF has been designed based on a block frequency domain adaptive filter using the well-known least-mean-square (LMS) criteria. There are two important coefficients in LMS adaptive filters which affect how an ESF adapts to changing acoustic conditions at each end of the conference, such as double-talk conditions and moving electroacoustic transducers. Previous approaches to similar ESFs have used either a single or double pair of these coefficients, whereas the new model typically uses ten. The performance of single, double, and multifilter architectures was compared. Performance was evaluated using both empirical measurements and subjective listening tests. Speech and music were used as the stimuli for a two-way teleconferencing experiment. The new filter performed better than the single- and two-filter ESF designs, especially in conferencing conditions with frequent double talk, and the new ESF can be optimized to suit different acoustic situations.

9:15

5aSP2. An adaptive equalization scheme using acoustic energy density. Xi Chen, Scott D. Sommerfeldt, and Timothy W. Leishman (ESC N283, Brigham Young Univ., Provo, UT 84602, drcxchen@hotmail.com)

This paper presents an equalization system based on the filtered-x algorithm and acoustic energy density. Previous theoretical results have shown that equalization filters using acoustic energy density produce greater spatial uniformity in equalization than those using acoustic pressure. The paper discusses the implementation of the equalization system in a one-dimensional duct. An adaptive algorithm is utilized in the time domain to determine and apply the equalization filter. Results from a traditional filter design and the energy density filter design are compared.

9:30

5aSP3. Separation and identification of a transient signal using the characteristics of eigenvectors. James P. Larue (AFRL/IFEC, Rome, NY 13441), George B. Smith (Naval Res. Lab., Stennis Space Ctr., MS 39529), George E. Ioup, and Juliette W. Ioup (Univ. of New Orleans, LA 70148)

This is an extension of previous research [Larue *et al.*, J. Acoust. Soc. Am. **113**, 2212 (2003)] into the physical characteristics possessed by eigenvectors in a singular value decomposition (SVD) of a covariance matrix formed from a sinusoidal signal corrupted by multipath to which Gaussian noise is added. An animation of the SVD in progress will focus on the creation of the eigenvectors rather than the singular values. In this case, the SVD forms a three-part subspace decomposition corresponding to the three components of the signal. These subspaces are clearly determined from characteristics (Fourier transform and the Kaiser Varimax norm) associated with the eigenvectors and are not so easily determined from the singular values alone. [Research supported by the NRC-AFRL/IFEC and ONR.]

9:45

5aSP4. Continuous speech recognition using dynamic synapse neural network. Alireza A. Dibazar, Hassan H. Namarvar, and Theodore W. Berger (BME Dept., Univ. of Southern California, 3650 S. McClintock Ave., OHE 500, CA 90089-1451)

The modified architecture of the dynamic synapse neural network (DSNN) is used to model windowed short time speech signal. The quasi-linearization algorithm is applied to train the network. The parameters of the trained network, which are representatives of the signal, are fed into the GMM/HMM based classifier. The performance of the modified architecture with GMM/HMM based classifier is demonstrated by recognition of continuous speech from unprocessed, noisy raw waveforms spoken by multiple speakers. Our results indicate that the features obtained from DSNN are robust in the presence of additive white Gaussian noise with respect to state-of-the-art Mel frequency features. [Work supported in part by DARPA, NASA and ONR.]

10:00–10:15 Break

10:15

5aSP5. A sound-texture detection algorithm. Michael J. Norris and Susan L. Denham (Ctr. for Theoretical and Computational Neurosci., Univ. of Plymouth, PL1 8AA, UK, mjnorris@plymouth.ac.uk)

A method for comparing sound textures based on the estimation of high-order information redundancy is presented. To calculate redundancy, a histogram is constructed that records the probability of occurrence of each possible short sequence of (typically about eight) samples. A self-organizing map [T. Kohonen, Proc. IEEE **78**, 1464–1480 (1990)] is used to summarize the histogram. The resulting probability distribution, or a trained self-organizing map that represents it, captures the texture of the sound, and leads to a natural definition of sound texture in terms of statistical stationarity. Sound textures such as recordings of running water, cafeteria noise, and traffic noise can be intuitively distinguished and classified by this method, despite small audible variations in amplitude, frequency, filtering, noise, and sound mixtures. The algorithm requires no preprocessing, Fourier analysis, heuristics, or finely tuned parameters, and may be implemented using a simple connectionist architecture. Textures can be compared through time by decaying probability distributions as new information arrives. A mechanism based on redundancy through time can also detect arbitrary changes in a sound, providing a novel mechanism to simulate auditory change detection. This versatile algorithm may find diverse applications in signal processing and neural computation. [Work supported by EU Open FET IST-2001-38099.]

10:30

5aSP6. Psychoacoustics based gain compensation for low listening levels. Thorvaldur Einarsson and Carol Espy-Wilson (Dept. of Elec. and Computer Eng., Univ. of Maryland, College Park, MD 20742)

Although the human hearing system is very complex and only understood to a limited degree, many measurements and models exist that explain parts of the hearing system. This paper uses one of these models, the contours of equal loudness, along with DSP techniques, to make music played at low listening levels sound more like it does at the intended listening level. The perceived frequency balance of music varies with listening level. This is especially noticeable at low listening levels, where frequencies below 500 Hz seem attenuated. Moreover, hearing perception exhibits nonlinear dynamic range compression, most evident at low frequencies. Traditional methods add fixed low-frequency gain to compensate for perceptual low-frequency attenuation at low levels. These methods do not consider the dynamic range compression and its nonlinearity and are often characterized by a boomy, unnatural sound. A system is designed where filter banks and power measurements estimate the time-varying power of low-frequency parts of the audio signal. The time-varying power of each narrow frequency band is compared to the contours of equal loudness, and changes are made to get the same frequency balance as at the intended listening level. The paper covers the design, implementation and performance of this system.

10:45

5aSP7. Two-dimensional Gabor analysis of space–time transient Lamb waves using Laser ultrasonic investigation. Loic Martinez (IUPGE ECIME UCP, 5 mail Gay Lussac, 95031 Cergy Pontoise, France, loic.martinez@iupge.u-cergy.fr), Nikolaas Van Riet, and Christ Glorieux (Katholieke Universiteit, Leuven, B-3001 Leuven, Belgium)

Laser generation/detection methods allow the investigation of ultrasonic transient phenomena in both space and time dimensions. Used for the experimental investigation of surface wave propagation along a one-dimensional medium, laser ultrasonics leads to two-dimensional (2D) space–time signal collections. In order to extract the wave propagation information, the classical high resolution signal processing methods or 2D Fourier transforms are not suitable to identify the transient and local aspect of wave propagation and mode conversion. In order to quantify these transient aspects in the space–time-wave number–frequency domains, the 2D Gabor transform is introduced. Its potential for the identification of the local and transient complex wave numbers is illustrated on the propagation of Lamb waves on a plane limited plate. The mode conversion sequences are clearly identified in the space–time–wave number-frequency planes. The experimental results are in good agreement whith the numerical simulations.

11:00

5aSP8. Phase velocimetry based on the spatio-temporal gradient analysis for detecting subsurface defects. Kenbu Teramoto (Dept. of Mech. Eng., SAGA Univ., Saga 8408502, Japan, tera@me.saga-u.ac.jp)

Ultrasonic Lamb-wave techniques are potential candidates of nondestructive evaluation (NDE) methodology: they allow the detection of surface defects or internal delamination. Because they propagate over a long distance, weakly attenuated in the case of a free-elastic plate, they are used for the evaluation of the assembly of large free plates in aeronautics, for example. The phase velocity of A0-mode Lamb-wave increases with thickness of the plate. Delaminations or internal defects, therefore, make the phase velocity slower. In this paper, an instantaneous phase velocimetry based on the spatio-temporal gradient analysis is proposed. The proposed method has an ability to measure the phase velocity of A0-mode Lamb wave field through the linearity among the four-dimensional vectors which is defined by following components: (1) a vertical displacement; (2) its vertical velocity; and (3) and (4) a pair of out-of-plane shear strains of the plate. In this study, the computational process of the local velocimetry is discussed and its physical meanings are investigated through FDTD simulations. Several results obtained through the proof-of-concept model are in relatively good agreement with the simulations.

11:15

5aSP9. Implementation of algorithms for extracting tonal components in underwater noise. M. H. Supriya and P. R. Saseendran Pillai (Dept. of Electron., Cochin Univ. of Sci. and Technol., Cochin-22, Kerala, India)

Underwater target classification and tracking problems utilize the noise signals emanating from the targets as well as the target dynamics features. Target classification is carried out by extracting certain classification clues about the target such as its emission frequencies, and other target specific features from the self noise or active transmissions. The self noise generated by the targets being nonstationary in nature are to be analyzed with short term averaged data segments for extracting the tonal components using various spectral estimation techniques. Though classical spectral analysis techniques are computationally efficient, it suffers from several inherent limitations such as frequency resolution, performance degradation due to implicit windowing of the data, etc. In an attempt to improve the spectral resolution, several modern spectral estimation techniques, which utilize the procedures of indirect Fourier analysis by fitting the measured short data segments to an assumed model, have been evolved. This paper presents the development of procedures for the estimation of power spectral densities using modern techniques based on parameter estimation such as auto regressive, moving average, auto regressive moving average, maximum likelihood, etc. The performance of the estimator has been validated by computing the emission frequencies of a 50-foot vessel.

Session 5aUW

Underwater Acoustics and Engineering Acoustics: Autonomous Underwater Vehicle Acoustics: Part I: Concepts and Systems

Henrik Schmidt, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Thomas R. Howarth, Cochair
NAVSEA Newport, 1176 Howell Street, Newport, Rhode Island 02841

Chair's Introduction—8:35

Invited Papers

8:40

5aUW1. Dolphin sonar detection and discrimination capabilities. Whitlow W. L. Au (Hawaii Inst. of Marine Biol., P.O. Box 1106, Kailua, HI 96734)

Dolphins have a very sophisticated short range sonar that surpasses all technological sonar in its capabilities to perform complex target discrimination and recognition tasks. The system that the U.S. Navy has for detecting mines buried under ocean sediment is one that uses Atlantic bottlenose dolphins. However, close examination of the dolphin sonar system will reveal that the dolphin acoustic hardware is fairly ordinary and not very special. The transmitted signals have peak-to-peak amplitudes as high as 225–228 dB *re* 1 μPa which translates to an rms value of approximately 210–213 dB. The transmit beamwidth is fairly broad at about 10o in both the horizontal and vertical planes and the receiving beamwidth is slightly broader by several degrees. The auditory filters are not very narrow with Q values of about 8.4. Despite these fairly ordinary features of the acoustic system, these animals still demonstrate very unusual and astonishing capabilities. Some of the capabilities of the dolphin sonar system will be presented and the reasons for their keen sonar capabilities will be discussed. Important features of their sonar include the broadband clicklike signals used, adaptive sonar search capabilities and large dynamic range of its auditory system.

9:00

5aUW2. AUVs as integrated, adaptive acoustic sensors for ocean exploration. Henrik Schmidt, Joseph R. Edwards, Te-Chih Liu, and Monica Montanari (Dept. of Ocean Eng., MIT, Cambridge, MA 02139, henrik@mit.edu)

Autonomous underwater vehicles (AUV) are rapidly being transitioned into operational systems for national defense, offshore exploration, and ocean science. AUVs provide excellent sensor platform control, allowing for, e.g., accurate acoustic mapping of seabeds not easily reached by conventional platforms, such as the deep ocean. However, the full potential of the robotic platforms is far from exhausted by such applications. Thus, for example, most seabed-mapping applications use imaging sonar technology, the data volume of which cannot be transmitted back to the operators in real time due to the severe bandwidth limitation of the acoustic communication. The sampling patterns are therefore in general being preprogramed and the data are being stored for postmission analysis. This procedure is therefore associated with indiscriminate distribution of the sampling throughout the area of interest, irrespective of whether features of interest are present or not. However, today's computing technology allows for a significant amount of signal processing and analysis to be performed on the platforms, where the results may then be used for real-time adaptive sampling to optimally concentrate the sampling in area of interest, and compress the results to a few parameters which may be transmitted back to the operators. Such adaptive sensing concepts combining environmental acoustics, signal processing, and robotics are currently being developed for concurrent detection, localization, and classification of buried objects, with application to littoral mine countermeasures, deep ocean seabed characterization, and marine archeology. [Work supported by ONR and NATO Undersea Research Center.]

9:20

5aUW3. New virtual sonar and wireless sensor system concepts. B. H. Houston, J. A. Bucaro, and A. J. Romano (Naval Res. Lab., 4555 Overlook Ave., SW Washington, DC 23075)

Recently, exciting new sensor array concepts have been proposed which, if realized, could revolutionize how we approach surface mounted acoustic sensor systems for underwater vehicles. Two such schemes are so-called ''virtual sonar'' which is formulated around Helmholtz integral processing and ''wireless'' systems which transfer sensor information through radiated RF signals. The ''virtual sonar'' concept provides an interesting framework through which to combat the dilatory effects of the structure on surface mounted sensor systems including structure-borne vibration and variations in structure-backing impedance. The ''wireless'' concept would eliminate the necessity of a complex wiring or fiber-optic external network while minimizing vehicle penetrations. Such systems,

however, would require a number of advances in sensor and RF waveguide technologies. In this presentation, we will discuss those sensor and sensor-related developments which are desired or required in order to make practical such new sensor system concepts, and we will present several underwater applications from the perspective of exploiting these new sonar concepts. [Work supported by ONR.]

9:40

5aUW4. Bluefin autonomous underwater vehicles: Programs, systems, and acoustic issues. Joseph E. Bondaryk (Bluefin Robotics Corp., 301 Massachusetts Ave., Cambridge, MA 02139, bondaryk@bluefinrobotics.com)

Bluefin Robotics Corporation has been manufacturing autonomous underwater vehicles (AUVs) since spinning out of the MIT Sea Grant Laboratory in 1997. Bluefin currently makes three different diameter models of AUVs; the 9, 12, and 21, all based on the same free-flooded architecture and vectored-thrust propulsion design. Auxiliary acoustic systems include acoustic abort, ranging beacons, and acoustic modems. Vehicle navigation is aided by a downward-looking acoustic Doppler velocity logger (DVL). Sonar payloads can include: bottom profiler, side-scan sonar, SAS, forward-looking imagers (DIDSON), as well as horizontal and vertical discrete hydrophone arrays. Acoustic issues that arise include: (1) transmission of sound through the ABS plastic vehicle shell; (2) the impact of vehicle self-noise on data; (3) interoperability of sonars with other acoustic emitters present on and off the vehicle; and (4) the impact of navigation on some acoustic operations like SAS. This talk will illustrate these issues with real data collected on various Bluefin vehicles.

10:00–10:20 Break

10:20

5aUW5. AUV-based synthetic aperture sonar: Initial experiences and insights. Daniel A. Cook, Jose E. Fernandez, John S. Stroud, Kerry W. Commander, and Anthony D. Matthews (Naval Surface Warfare Ctr., Panama City, Code R21, Panama City, FL 32407-7001)

The ability to do synthetic aperture sonar (SAS) imaging from autonomous underwater vehicles (AUVs) has only recently been achieved. The combination of the two technologies is a milestone in the field of underwater sensing as the combination of high-resolution SAS imaging with AUVs will provide military, research, and commercial users with systems of unprecedented performance and capabilities. The U.S. Navy took delivery of the first AUV-based SAS in early 2003. A description of the system will be presented along with the methodologies employed. The emphasis will be on image quality and repeatability, as well as the differences associated with operating an AUV-based SAS as opposed to a towed SAS. Additional topics will include general comments and recommendations for better AUV/SAS integration such as mission planning and vehicle control strategies intended to maximize the chances of high-quality imagery, SAS motion measurement requirements coupled with on-board vehicle navigation, and the potential of using the SAS data to augment the vehicle motion sensors. Lastly, a brief overview of forthcoming Navy SAS systems will be included.

10:40

5aUW6. Passive acoustic localization with an AUV-mounted hydrophone array. Gerald L. D'Spain, Eric Terrill, C. David Chadwell, Jerome A. Smith, and Richard Zimmerman (Marine Phys. Lab., Scripps Inst. of Oceanogr., La Jolla, CA 93940-0704)

A mid-size Odyssey IIb autonomous underwater vehicle (AUV) was retrofitted with the advanced vectored-thrust system presently installed on AUVs manufactured by Bluefin Robotics, Inc. Subsequent modifications to this thrust system decreased the radiated acoustic and vibration noise levels recorded by an eight-element hydrophone array mounted on the AUV's inner shroud by 20 to 50 dB across the 20 Hz to 10 kHz band. This reduction in self-noise levels to near, or at, background ocean noise levels permits the use of the vehicle-mounted hydrophone array in passive ocean acoustic studies. One example is the application of passive synthetic aperture processing techniques to provide greater spatial resolution estimates of the direction of low frequency sources. Doppler spreading caused by medium motion is a limiting factor in array gain. At mid frequencies (1–10 kHz), the complexity of the received acoustic field created by scattering off the AUV body is partly captured in the array processing by the use of replica vectors measured in a calibration tank. These empirical replica vectors decrease the azimuthally dependent degradation in beamforming performance over that of plane waves. [Work supported by ONR, Code 321(US).]

Contributed Papers

11:00

5aUW7. Exploiting autonomous underwater vehicle mobility for enhanced sonar performance. Joseph R. Edwards and Henrik Schmidt (MIT Dept. of Ocean Eng., 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139)

Autonomous underwater vehicles (AUVs) provide a mobile sonar platform for local environmental characterization and mine hunting missions in areas that are prohibitively dangerous or otherwise inaccessible to manned vessels. Such missions are typically implemented by providing the vehicle with a series of way-points that uniformly sample the region of interest. While these preplanned sampling paths can be effective for many purposes, a higher degree of efficiency and efficacy can be achieved when the AUV is able to react to its perceived environment. Such response to the sensory stimulus represents a preliminary step toward the AUV achieving tasks in the way that a dolphin or human might. Upon sensing an object or feature of interest, the AUV can further interrogate the object by obtaining multiple views from preferred vantage points. Probabilities of detection and correct classification can be greatly impacted by a wise choice of sonar-adaptive AUV behavior. In this paper, the ability of the AUV to improve sonar performance through sonar-adaptive mission planning is investigated in the context of the mine hunting problem. Trade-offs

are discussed for continuous versus scripted mission adaptations, single-minded versus multiple-objective mission control, and task-wise versus mission-wise behavior. [Work supported by ONR and SACLANTCEN.]

11:15

5aUW8. Surface ship wake remote sensing using an AUV-mounted multibeam sonar. Lee Culver and David Bradley (Appl. Res. Lab., The Penn State Univ., P.O. Box 30, State College, PA 16804)

A 250-kHz multibeam sonar integrated into a high-speed autonomous underwater vehicle (AUV) has been used to image the wakes of two large surface ships from the underside. Using vehicle autopilot data (yaw, pitch, and roll) and acoustic tracking data, the multibeam sonar data were processed to produce 3D maps of the ship wake intensity. Also, the backscattered signal has been used to estimate the number of bubbles per unit volume (the bubble density), assuming a particular size distribution of the bubbles, using a technique that other researchers have employed to estimate the bubble density from single beam sonar data. These measurements have been discussed in two earlier ASA talks [J. Acoust. Soc. Am. **110**, Pt. 2 (2001); **111**, Pt. 2 (2002)]. The present talk focuses on how vehicle motion is used to process acoustic data from individual beams so as to produce a 3D map of the ship wake. Future plans involving a commercial multibeam sonar integrated into a high speed vehicle are discussed. [Work sponsored by NAVSEA PMS 415 and ONR Code 333.]

11:30

5aUW9. Concurrent detection and classification of targets with multistage signal-processing algorithms. Monica Montanari, Joseph R. Edwards, and Henrik Schmidt (Dept of Ocean Eng., MIT, 77 Massachusetts Ave., Rm. 5-204, Cambridge, MA 02139, momo@mit.edu)

Concurrent detection and classification (CDAC) of targets stands as the goal in littoral mine-hunting missions. CDAC systems commonly apply model-based algorithms that include *a priori* known features of the target inside the detection algorithm. If the models are accurate, then this approach significantly reduces the false-alarm rate inherent in detection-only methods. When the possible targets are unknown, as may be the case in tactical situations, then these model-based methods not only fail to reduce the false-alarm rate, but may also reduce the probability of detection. Simultaneous optimization of detection and classification presents a challenge due to competing criteria; detection seeks to integrate signals to improve signal-to-noise ratio, while classification seeks to preserve small features of distinction within the signals. In this work, a method for robust CDAC is demonstrated that exploits the capabilities of autonomous underwater vehicles (AUVs) and multistage signal-processing algorithms to systematically investigate targets of interest in a single mission. The deformable geometry of the AUV-borne sonar network is exploited to provide favorable views of targets to achieve multiple objectives in series, and on-board computational facilities allow the implementation of multiple signal-processing regimes [Work supported by ONR and NATO Undersea Research Centre.]

11:45

5aUW10. Intervessel navigation using range and range rate. Brian S. Bourgeois and Patrick M. McDowell (Naval Res. Lab., Stennis Space Ctr., MS 39529, bsb@nrlssc.navy.mil)

A fundamental requirement for groups of unmanned underwater vehicles to work cooperatively together is the ability for each vessel to know the relative position of its neighbors. While external communication and positioning infrastructures can be used for this purpose, a more flexible approach is to give each vessel the ability to independently discern the location of its neighbors using its own communication and sensor systems. Recent developments in acoustic modems have included the ability to measure range between two vessels as well as the Doppler imparted on the signal by the relative motion between vessels. The measured Doppler shift can be used to compute the instantaneous range rate between two vessels; researchers at NRL are presently working to validate these measurements using GPS. Approaches to vessel relative positioning and navigation using successive range and range-rate measurements, along with directed vessel maneuvers, will be presented.

FRIDAY AFTERNOON, 28 MAY 2004 NEW YORK BALLROOM A, 1:00 TO 2:30 P.M.

Session 5pAOa

Acoustical Oceanography: Ocean Basin Acoustics

John L. Spiesberger, Chair

Department of Earth and Environmental Science, University of Pennsylvania, 158 Hayden Hall, 240 South 33rd Street, Philadelphia, Pennsylvania 19104-6316

Contributed Papers

1:00

5pAOa1. Acoustic identification of a single transmission at 3115 km from a bottom-mounted source at Kauai. John L. Spiesberger (Dept. of Earth and Environ. Sci., 240 S. 33rd St., Univ. of Pennsylvania, Philadelphia, PA 19104, johnsr@sas.upenn.edu)

Sounds received in the Gulf of Alaska at 3115 km from the ATOC/NPAL source at Kauai (75-Hz, 0.027-s resolution, bottom-mounted) are compared with acoustic and oceanographic models. Unlike data collected at stationary SOSUS arrays, these data come from a towed horizontal array at 372-m depth of military origin. A plausible identification of the acoustic reception is made despite the fact that only one transmission is collected and sound interacts with the bottom near the source. The similarity between the modeled and measured impulse response here may be useful for understanding the signals between this same source and the NPAL array near southern California. The plausible identification of sound from the horizontal array here appears to point toward the feasibility of using other military platforms of opportunity besides SOSUS to study acoustic propagation and possibly map climatic changes in temperature by means of tomography.

1:15

5pAOa2. U.S. Navy sources and receivers for studying acoustic propagation and climate change in the ocean. John L. Spiesberger (Dept. of Earth and Environ. Sci., 240 S. 33rd St., Univ. of Pennsylvania, Philadelphia, PA 19104, johnsr@sas.upenn.edu)

Sounds from a U.S. Navy SSQ-110A source are received at high signal-to-noise ratios at ocean-basin scales at two Sound Surveillance Systems in the Pacific. The sounds have sufficient pulse resolution to study climatic variations of temperature. The acoustic data can be understood

using ray and parabolic approximations to the wave equation. Modeled internal waves decrease pulse resolution from 0.01 to 0.1 s, consistent with observations.

1:30

5pAOa3. Science enabled by ocean observatory acoustics: The NSF ORION program. Bruce M. Howe (Appl. Phys. Lab., Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698, howe@apl.washington.edu), James H. Miller (Univ. of Rhode Island, Narragansett, RI 02882), and IASOO Committee (Integrated Acoust. Systems for Ocean Observatories Committee, Acoust. Society of America)

The National Science Foundation (NSF) has started the Ocean Research Interactive Observatories Network (ORION) program for research-driven sustained observations. The core infrastructure will consist of: (1) a coarse global array of buoys, (2) a regional cabled observatory in the northeast Pacific (with Canada already funded for the northern portion), and (3) coastal observatories. Seafloor junction boxes providing power and communications are a common enabling feature. The ORION Workshop (Puerto Rico, 4–8 January 2004) developed science themes that can be addressed utilizing this infrastructure. Acoustics enable much of the science. The use of acoustics to sense the synoptic 3-D/volumetric ocean environment was found to be ubiquitous through most ORION working groups. One reason for this is the relative transparency of the ocean to sound and the opaqueness to electromagnetic radiation. Participants at the workshop formed an Acoustics Working Group. Based on its report, we review the science and technical drivers for acoustics and educational opportunities. Themes include inherent volumetric, near instantaneous sampling, robust transducers, imaging at many scales, navigation, communications, and using sound in the sea as a major education and outreach mechanism. Recommendations include the formation of a standing ORION committee on acoustics and a workshop. See http://www.orionprogram.org and http://www.oce.uri.edu/ao/AOWEBPAGE.

1:45

5pAOa4. Horizontal coherence in the NPAL experiment. Michael Vera, Matthew Dzieciuch (Scripps Inst. of Oceanogr., Univ. of California, San Diego, La Jolla, CA 92093-0225), and The NPAL Group[a] (APL-UW, SIO-UCSD, WHOI)

Acoustic transmissions from a broadband source near Kauai with a center frequency of 75 Hz were recorded on a two-dimensional receiver array at a range of 3900 km as part of the North Pacific Acoustic Laboratory (NPAL) experiment. The receiver array consisted of five vertical line arrays (VLAs), with separations transverse to the propagation path ranging from a few hundred meters to a few kilometers. The coherences derived from the data have been compared to two different numerical predictions based on sound-speed perturbations due to internal waves. An approximation to the acoustic path integral yields a prediction for the length scale of horizontal coherence. Parabolic-equation simulations of propagation from the source to each VLA through multiple realizations of a random internal-wave field provide another estimate of the coherence at each horizontal separation. [Work supported by ONR.] [a]J. A. Colosi, B. D. Cornuelle, B. D. Dushaw, M. A. Dzieciuch, B. M. Howe, J. A. Mercer, R. C. Spindel, and P. F. Worcester.

2:00

5pAOa5. Precise measurement of travel time in 1999 OAT experiment. Wang Yong (Grad. of Sci. & Technol., Chiba Univ., 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan, wang@graduate.chiba-u.jp) and Hachiya Hiroyuki (Chiba Univ., Chiba 263-8522, Japan)

The 1999 OAT experiment is performed in the Central Equatorial Pacific to monitor the ocean phenomenon associated with El Niño and the Southern Oscillation (ENSO). In order to analyze the travel-time perturbations due to ocean current, amplitude information of the received signal is used in the conventional method. However, it is difficult to measure the travel-time difference precisely, mainly because of the attenuation of long-range transmission. In this report, we propose a new technique which is the complex vector method by using the amplitude and phase information. From the ray identification results, we estimated the travel-time difference between corresponding rays of the reciprocal transmissions. Since the phase of a ray signal is very stable in 130-s signal duration time, the precision of the measurement of the travel-time difference using phase difference is higher than the conventional estimation. The travel-time differences estimated by using this method have reasonable magnitudes.

2:15

5pAOa6. Sensitivity kernels of finite-frequency travel times in ocean acoustic tomography. Emmanuel K. Skarsoulis (Inst. of Appl. and Comput. Math. FORTH, 711 10 Heraklion, Crete, Greece, eskars@iacm.forth.gr) and Bruce D. Cornuelle (Scripps Inst. of Oceanogr., La Jolla, CA 92093)

Wave theoretic modeling is applied to obtain travel-time sensitivity kernels representing the amount by which travel times are affected by localized sound-speed variations anywhere in the medium. In the ray approximation travel times are sensitive to medium changes only along the corresponding eigenrays. In the wave-theoretic approach the perturbations of peak arrival times are expressed in terms of pressure perturbations, which are further related with the underlying sound-speed perturbations using the first Born approximation. In this way, an integral representation of travel-time perturbations is obtained in terms of sound-speed perturbations; the associated kernel represents the spatial sensitivity of travel times to sound-speed perturbations. The application of the travel-time sensitivity kernel to an ocean acoustic waveguide gives a picture close to the ray-theoretic one in the high-frequency case but significantly differs at lower frequencies. Low-frequency travel times are sensitive to sound-speed changes in areas surrounding the eigenrays, but not on the eigenrays themselves, where the sensitivity is zero. Further, there are areas of positive sensitivity, where, e.g., a sound-speed increase results in a counter-intuitive increase of arrival times. These findings are confirmed by independent forward calculations.

Session 5pAOb

Acoustical Oceanography: Passive and Active Observations

Nicholas C. Makris, Chair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Contributed Papers

2:45

5pAOb1. Prediction of ocean ambient sound generated from geophysical signals. Barry Ma, Jeffrey Nystuen, and Ren-Chieh Lien (APL, Univ. of Washington, 1013 NE 40th St., Seattle, WA 98105-6698)

About 90 buoy months of ocean ambient sound data are collected using Acoustic Rain Gauges in different open-ocean locations. Distinct ambient sound spectra are identified through a series of discrimination process. Some distinct features are described as follows: (1) Excluding effects of rain, observed ambient sound spectra, generated exclusively by wind, exhibit a constant frequency spectral slope in 1–35 kHz. (2) Drizzle produces a prominent spectral peak at 15 kHz, and the magnitude of the spectral peak is very sensitive to the wind speed. These are consistent with earlier findings [Vagle *et al.* (1990) and Nystuen (1993)]. (3) Preliminary analysis shows that the spectral slope between 1–10 kHz decreases linearly with the rainfall rate. Therefore, the rainfall rate may be reliably predicted by both the spectral slope and the spectral magnitude in 1–10 kHz frequency range. The comparison of these two independent predictions provides the quality check of the estimate of the rainfall rate. An analytical spectral model of the ocean ambient sound from 1–50 kHz is constructed using two input parameters, rainfall rate and wind speed, based on the existing observations. This analytical model is aimed at predicting the ambient sound spectra at varying rainfall rates and wind speeds.

3:00

5pAOb2. A novel technique for measuring the rainfall from the rain noise. Mani Thundiyil and P. R. Saseendran Pillai (Dept. of Electron., CUSAT, Cochin 682022, India)

It has been observed that if water droplets are allowed to fall with terminal velocity, on to the water trapped in the chamber of a specially designed sensor assembly, the resulting acoustic signal generated from the drop impact is wideband and extends up to 100 kHz. The acoustic energy of the signal in the low-frequency range is seen to be vis-a-vis correlated to the kinetic energy of the drops. Such an assembly is exposed to the rain to pick up the rain-generated noise. The energy of the rain-generated noise is seen to be proportional to the kinetic energy of the rain. From the rain kinetic energy–intensity relationship, the rainfall rate is computed. This paper presents an experimental procedure for estimating the rain intensity from the rain generated underwater noise captured by the sensor assembly. Experiment has been carried out for various rainfall rates and the computed rain intensities were compared with the ones measured with a tipping bucket rain gauge. The results and inferences suggest that this approach is a simple, cost effective and computationally efficient technique to measure the kinetic energy of rain as well as compute the rainfall rate.

3:15

5pAOb3. Experimental demonstration of accurate hurricane classification from local wind speed estimates obtained with underwater sound. Joshua D. Wilson and Nicholas C. Makris (77 Massachusetts Ave. 5-212, Cambridge, MA 02139, makris@mit.edu)

In 1999 Hurricane Gert passed over an autonomous hydrophone in the North Atlantic yielding a clear recording of the underwater noise generated by the hurricane. By correlating the noise with meteorological data from reconnaissance aircraft and satellites, it is shown that underwater noise intensity in the 10–50-Hz band is proportional to local wind speed to the 3.3 power. Previous experiments have shown similar relationships at lower wind speeds in many different ocean environments [Crouch, J. Acoust. Soc. Am. **51**, 1066–1072 (1972); Marrett and Chapman, IEEE J. Ocean. Eng. **15**, 311–315 (1990); Chapman and Cornish, J. Acoust. Soc. Am. **93**, 782–789 (1993)]. Given this empirical relationship, it is shown that acoustic intensity measurements can be used to accurately estimate wind speed to within a 5% error margin, and so can be used to accurately classify the destructive power of a hurricane. Hurricane Gert, for example, is found to be a class 2 hurricane with maximum wind speeds of at least 89 kts by this underwater acoustic approach, which is in accord with aircraft measurements. Potential advantages of this underwater acoustic approach are discussed as are their potential impact on hurricane forecasting and disaster planning.

3:30

5pAOb4. Explaining extended linear features observed in remote sonar images of the New Jersey continental shelf break during Acoustic Clutter Experiments in 2001 and 2003. Sunwoong Lee, Purnima Ratilal, and Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

Prominent acoustic clutter forming a lineated feature spanning more than 10 km has been repeatedly observed in long-range sonar imagery acquired during the Acoustic Clutter Experiments in 2001 and 2003. The lineated feature appears on the continental shelf near the shelf break for downslope sonar positions. It is hypothesized that this lineated feature is caused by scattering from fish shoals, seafloor slopes, or subbottom geology such as the R-reflector. A range-dependent reverberation and scattering model using the parabolic equation [Ratilal *et al.*, J. Acoust. Soc. Am. **114**, 2302 (2003)] is employed to test these various hypotheses by model–data comparison using high resolution bathymetry, subbottom profiles, and fish density data acquired in support of the long-range acoustic measurements.

3:45

5pAOb5. Inferring fish school distributions from long range acoustic images: Main acoustic clutter experiment 2003. Deanelle T. Symonds, Purnima Ratilal, Nicholas C. Makris (MIT, 77 Massachusetts Ave., Cambridge, MA 02139), and Redwood W. Nero (Naval Res. Lab, Stennis Space Ctr., MS 39529)

Long range scattering from fish schools and bottom reverberation in the New Jersey Continental Shelf environment are modeled using a unified, range-dependent, bistatic scattering, and reverberation model based on the parabolic equation [Ratilal and Makris, J. Acoust. Soc. Am. **114**, 2302 (2003)]. The fish swim bladder is approximated as an air-filled bubble, while the bottom reverberation from volume inhomogeneities is modeled using the Rayleigh–Born approximation. The broadband scattered field, in the frequency range from 390 to 440 Hz, is beamformed and spatially charted using two-way travel time. The model output is compared with scattered field levels from fish schools and background reverberation measured during the the Main Acoustic Clutter Experiment 2003 using a long range, bistatic sonar system. The fish school characteristics, such as size, distribution and density, are inputs to the model. These are obtained

from measurements made by the fish finding sonar during the experiment. This calibrated model is then used to infer fish school distributions and densities in areas where fish finding sonar measurements are not available.

4:00

5pAOb6. Continuous wide area monitoring of fish shoaling behavior with acoustic waveguide sensing and bioclutter implications. Nicholas C. Makris, Purnima Ratilal, Deanelle T. Symonds (MIT, 77 Massachusetts Ave., Cambridge, MA 02139, makris@mit.edu), and Redwood W. Nero (Naval Res. Lab, Stennis Space Ctr., MS 39529)

Field measurements are used to show that the detailed behavior of fish shoals can be continuously monitored at roughly 1-min intervals over wide areas spanning hundreds of square kilometers by long range acoustic waveguide sensing. The technique was used on the New Jersey Continental Shelf to produce unprecedented video images of shoal formation, fragmentation, and migration. Simultaneous line-transect measurements show the imaged shoals to contain pelagic fish with densities of at least one individual per meter^3. The technique relies upon acoustic waveguide propagation in the continental shelf. Here, trapped modes dominate propagation and suffer only cylindrical spreading loss rather than the spherical loss suffered in free-space transmission or short-range propagation in the ocean. In contrast, standard methods for fish surveyance involve line transect measurements from slow moving research vessels that significantly under-sample fish distributions in time and space, leaving an incomplete behavioral picture. The implications of this bioclutter phenomenon on the Navy's long range active sonar operations in continental shelf environments are discussed.

FRIDAY AFTERNOON, 28 MAY 2004 NEW YORK BALLROOM B, 1:25 TO 5:15 P.M.

Session 5pEA

Engineering Acoustics and Underwater Acoustics: Autonomous Underwater Vehicle Acoustics: Part 2: Hardware and Devices

Thomas R. Howarth, Cochair
NAVSEA Newport, 1176 Howell Street, Newport, Rhode Island 02841

Henrik Schmidt, Cochair
Department of Ocean Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139

Chair's Introduction—1:25

Invited Papers

1:30

5pEA1. Broadband, lower-frequency acoustic projector design for AUV applications. Thomas R. Howarth (U.S. Navy, Newport, RI and Washington, DC, thowarth@ccs.nrl.navy.mil), Kim C. Benjamin, Dehua Huang (Naval Sea Systems Command, Div. Newport, Newport, RI), James F. Tressler (Naval Res. Lab., Washington, DC), and Walter Carney (Naval Sea Systems Command, Crane, IN)

Autonomous underwater vehicles (AUVs) offer a difficult platform for mounting lower-frequency acoustic projectors. Researchers at various U.S. Navy laboratories have been coordinating their interests to address the concepts and reduction to practice of lower frequency and broader operating frequency bands transduction technologies. Recent efforts have resulted in the development of cymbal-based acoustic transmitting devices to operate over an extended frequency range of 1 to 25 kHz within a package that can be housed inside of a BlueFin AUV. This presentation will discuss the concepts, design, development, and integration of the cymbal-based modular panels and how they are being configured to meet these applications. *In situ* data, including in-water calibration of acoustic performance, will be provided. Discussions for ongoing work and future directions will conclude the presentation. [Work supported by the ONR.]

1:50

5pEA2. 1-3 Piezocomposite transducers for AUV applications. Brian Pazol, Ken Lannaman, and Barry Doust (Mater. Syst., Inc., 543 Great Rd., Littleton, MA 01460)

Sonar systems on board AUVs present interesting challenges to the transducer designer because of their small size, low weight requirements, and limited available power. 1-3 piezocomposite transducers offer many performance characteristics which make them ideal for deployment in AUVs. Piezocomposite transducers are light weight, have broad bandwidth, have high efficiency, and can be conformed to fit the curvature of the vehicle. The broad bandwidths and low sidelobes made possible by piezocomposites result in sharper images with less distortion. The piezocomposite material is mechanically robust and can survive the rigors of normal operations as well as AUV deployment and retrieval. In addition, the conformal configuration substantially reduces hydrodynamic drag. As a conformal array, there is nothing to get knocked off during deployment and retrieval operations, or entangled with natural or man-made objects suspended in the water column. This contributes directly to improving the operational endurance of the AUV system, thereby enhancing overall system utility. MSI has produced and tested a variety of piezocomposite transducers for use in obstacle avoidance, mine hunting, and acoustic communications. An overview of piezocomposites and recent results of piezocomposite transducers will be presented.

2:10

5pEA3. Acoustic communications and autonomous underwater vehicles. Lee Freitag, Matthew Grund, James Preisig (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543, lfreitag@whoi.edu), and Milica Stojanovic (MIT, Cambridge, MA 02139)

Acoustic communications systems used on autonomous underwater vehicles (AUVs) provide supervisory control, access to real-time data, and also allow multiple vehicles to cooperate in undertaking adaptive sampling missions. However, the use of acoustic systems on AUVs presents special challenges because of limited space for optimal placement of transducers, and potential conflicts with other acoustic systems such as side-scan sonars and transponders. In addition, radiated and structure-borne acoustic interference from thrusters and actuators reduces the sensitivity of on-board receivers. Recent work in acoustic communications and AUVs has included combining some navigation functions into communications equipment, development of operating modes that remove conflicts between different subsystems, design of vehicle components to avoid or remove interference, and other approaches to improving performance. While these efforts have been successful for specific installations, many challenges remain. This talk addresses problems and solutions for supervised and completely autonomous multi-vehicle communications to support complex AUV missions. Also presented are recent results which demonstrate that acoustic communications can be used successfully on a variety of AUV platforms for many different applications. [Work supported by ONR.]

2:30

5pEA4. Acoustic pressure-vector sensor array. Dehua Huang, Roy C. Elswick (NUWC, Newport, RI 02841), and James F. McEachern (ONR, Arlington, VA 22217)

Pressure-vector sensors measure both scalar and vector components of the acoustic field. December 2003 measurements at the NUWC Seneca Lake test facility verify previous observations that acoustic ambient noise spectrum levels measured by acoustic intensity sensors are reduced relative to either acoustic pressure or acoustic vector sensor spectrum levels. The Seneca measurements indicate a reduction by as much as 15 dB at the upper measurement frequency of 2500 Hz. A nonlinear array synthesis theory for pressure-vector sensors will be introduced that allows smaller apertures to achieve narrow beams. The significantly reduced ambient noise of individual pressure-vector elements observed in the ocean by others, and now at Seneca Lake, should allow a nonlinearly combined array to detect significantly lower levels than has been observed in previous multiplicative processing of pressure sensors alone. Nonlinear array synthesis of pressure-vector sensors differs from conventional super-directive algorithms that linearly combine pressure elements with positive and negative weights, thereby reducing the sensitivity of conventional super-directive arrays. The much smaller aperture of acoustic pressure-vector sensor arrays will be attractive for acoustic systems on underwater vehicles, as well as for other applications that require narrow beam acoustic receivers. [The authors gratefully acknowledge the support of ONR and NUWC.]

2:50

5pEA5. Advanced sonar array concepts for small underwater vehicle applications. Kim C. Benjamin (NAVSEA Div. Newport, 1176 Howell St., Newport, RI 02841)

Low-frequency acoustic reception and high directivity are difficult to obtain simultaneously using small undersea (robot) platforms. The dimensions and geometry of the hydrodynamic shapes that are required for efficient transit through the water preclude the use of broad apertures required for low-frequency directional applications. This paper will present array concepts that could be considered the next step in developing low-frequency directional acoustic payloads for undersea research and exploration. The talk will include results of previous Navy experimental demonstrations of wide aperture arrays for small vehicles as well as future advanced sonar packaging concepts that rely on new transduction technologies. [Work supported by the U.S. Navy.]

3:10–3:30 Break

Contributed Papers

3:30

5pEA6. Compact, high power, energy efficient transmit systems for UUVs using single crystal transducers. Harold Robinson (NUWC Div. Newport, Code 2132, Bldg. 1170, 1176 Howell St., Newport, RI)

UUVs are currently being designed to perform a multiplitude of tasks in ocean exploration and Naval warfighting. Many of these tasks require the use of active acoustic projectors, and many may require the UUV to operate independently for hours, days, or even weeks. In order for a UUV to be as versatile as possible, its active transmit system must be versatile as well, implying that broad acoustic bandwidths are a must. However, due to size and battery life limitations, this broadband system must also be compact and energy efficient. By virtue of their extraordinary material properties, ferroelectric single crystals are the ideal transduction material for developing such broadband systems. The effect of their high coupling factor on transmit systems shall be illustrated by showing the dramatic impact on amplifier size, power factor, and acoustic response that is possible using these materials. In particular, a transducer built with these materials can be well matched to the power amplifier, i.e., 80% or more of the amplifier power reaches the transducer, over decades of frequency. Measured results from several prototype single crystal transducers shall be presented to demonstrate that the theoretical gains are actually realizable in practical devices. [Work sponsored by DARPA.]

3:45

5pEA7. A dual-piston ring-driven X-spring transducer. Alexander L. Butler, John L. Butler (Image Acoust., Inc., 97 Elm St., Cohasset, MA 02025, abutler@imageacoustics.com), Robert L. Pendleton, and Richard M. Ead (Naval Undersea Warfare Ctr., Newport, RI 02841)

Tonpilz transducers generally consist of a stack of piezoelectric material sandwiched between a single piston and an inertial tail mass or between two pistons. The result is a transducer with a large length-to-diameter ratio. The X-spring transducer design, based on U.S. Patent 4 845 688, allows a means for a shorter transducer length through an orthogonal piezoelectric drive system coupled to the pistons by lever arms. We present here a low-frequency, dual-piston piezoelectric ceramic ring driven version with a length of only 10 in. and a diameter of 19 in. Both single-element and two-element array results are presented. The measured response is shown to be in agreement with the finite-element model with a

smooth, wideband 300- to 550-Hz response for this dual-piston, ring-driven X-spring transducer. [Work supported by a Phase II SBIR, through NUWC, Newport, RI 02841.]

4:00

5pEA8. Acoustic cymbal performance under hydrostatic pressure. Kirk E. Jenne, Dehua Huang, and Thomas R. Howarth (NAVSEA Div. Newport, 1176 Howell St., Newport, RI 02841-1708, JenneKE@npt.nuwc.navy.mil)

Continual awareness about the need to develop light-weight, low-volume, broadband, underwater acoustic projector and receive arrays that perform consistently in diverse environments is evident in recent Navy acoustic system initiatives. Acoustic cymbals, so named for resemblance to the percussive musical instruments, are miniature flextensional transducers that may perhaps meet the performance criteria for consistent performance under hydrostatic pressure after modifications in the design. These acoustic cymbals consist of a piezoceramic disk (or ring) bonded to two opposing cymbal-shaped metal shells. Operating as mechanical transformers, the two metal shells convert the large generative force inherently within the disk's radial mode into increased volume displacement at the metal shell surface to obtain volume displacement that translates into usable source levels and/or sensitivities at sonar frequencies in a relatively broad band. The air-backed design for standard acoustic cymbal transducers presents a barrier to deepwater applications. A new acoustic cymbal design for high-pressure applications will be presented for the first time. This practical pressure compensation is designed to diminish the effects of hydrostatic pressure to maintain consistent acoustic cymbal performance. Transmit and receive performance data, determined at the Naval Undersea Warfare Center's (NUWC) Acoustic Pressure Tank Facility (APTF), is presented.

4:15

5pEA9. An acoustic array to be towed behind an autonomous underwater vehicle. Jason D. Holmes, William M. Carey (Boston Univ., Boston, MA 02215, jholmes@bu.edu), and James F. Lynch (Woods Hole Oceanogr. Inst., Woods Hole, MA 02543)

The use of a towed array from a small autonomous underwater vehicle, AUV, such as the WHOI-Remus vehicle is discussed as a valuable ocean acoustics measurement tool for of 3-D characterization of shallow water regions. The feasibility of towing a 10-m-long, small-diameter fluid-filled hydrophone array behind Remus was investigated with a laboratory-prototype array and preliminary tests from the WHOI pier. Preliminary results on array self-noise are presented for low tow speeds. The flow noise, vehicle noise, and other unwanted signal degrading noise sources are examined, discussed, and shown by theoretical arguments to be reduced to low levels by proper array design. This paper examines current technology for such an array with digital sampling and recording equipment incorporated in the Remus vehicle. Consideration is also given to the deployment problem of array on an autonomous vehicle. [Work supported by the College of Engineering, Boston University.]

4:30

5pEA10. A framework of concurrent navigation and seabottom targets detection using acoustic sensors on AUVs. Te-Chih Liu and Henrik Schmidt (MIT, 77 Massachusetts Ave., Cambridge, MA 02139)

The use of Autonomous Underwater Vehicles (AUVs) for Mine Counter Measures (MCMs) is an area of active recent research. The excellent mobility of AUVs allows for multi-aspect sonar view of the targets for improved detection, tracking, and classification. However, the uncertainty of the platforms and associated target localization degrade the detection ability by AUVs. Furthermore, the weak signals from buried targets is another severe problem making detection by a single measurement impossible. A new acoustic sensing framework is proposed for detection and tracking seabottom targets based on the track-before-detect (TBD) technique. In contrast to the traditional methods, TBD tracks possible targets before the detection is declared. There are several advantages by using this framework: (1) Compared to the traditional method of detection followed by tracking, higher detection probability is achieved for dim target detection due to the integrated detection metric of TBD. (2) Compared to the multiple hypothesis tracker (MHT) method, instead of searching a diverge hypotheses tree, TBD speeds up the searching and decreases the computational load, which makes onboard implementation feasible. (3) The stochastic models of uncertainties of targets and AUVs are based on the Bayesian framework, and thus, it is easy to apply various recursive estimators such as the Kalman filter or particle filter for tracking individual targets as well as the AUV platforms. Results of a successful application of this method in the GOATS2002 experiment are demonstrated. [Work supported by ONR and NATO Undersea Research Centre.]

4:45

5pEA11. Acoustic navigation for autonomous underwater vehicles. Wen-Bin Yang and T. C. Yang (Naval Res. Lab., Code 7120, 4555 Overlook Ave. SW, Washington, DC 20375, wyang@wave.nrl.navy.mil)

To effectively use groups of autonomous underwater vehicles (AUVs), an accurate navigation capability must be developed. We describe an acoustic navigation technique that uses an acoustic communications probe signal to estimate the range, range rate and direction of arrival signal. The acoustic signal may be transmitted from one vehicle or node (e.g., AUV, ship or fixed mooring) to another. The technique was tested during the RAGS03 experiment. A towed acoustic communications receiver that replicated the motion of an AUV and a fixed source node were used. The receiver was navigated with a global position system (GPS). The range, range rate and direction of arrival signal have been estimated using the acoustic signal received on a horizontal line array. The results are compared to the known GPS source and receiver positions. The navigation technique will be discussed. Its advantages over the traditional navigation by multiple transponders will be outlined. [Work supported by the ONR.]

5:00

5pEA12. Experimental modeling of the acoustic signature of an AUV. Joseph M. Cuschieri (Lockheed Martin Corp., MS2 Undersea Systems, Riviera Beach, FL 33404)

In this presentation test tank results from an experimental analysis of the acoustic signature of an Ocean Explorer class AUV are presented. The results are from measurements performed in a reverberant test tank on an AUV model, and of an AUV under typical operating conditions. The main source of excitation is the propulsion module or ''podule.'' Inside the podule are the propulsion motor and the motors for the control surfaces with penetrations for the main propulsion shaft and the shafts for the control surfaces. Different operating conditions and different mountings for the main propulsion and control module (podule) inside the AUV are considered. The influence of the propeller, the mounting of the podule, and covering the podule with a compliant layer are considered in the measurements. It is shown that for the type of AUV considered here, the mounting of the podule is not very significant and that significant energy is transferred through the water trapped between the podule and the AUV hull. Furthermore, the propeller has a significant influence on the acoustic signature. [Work supported by ONR.]

Session 5pNS

Noise: Noise in Large Cities II

Daniel R. Raichel, Chair
2727 Moore Lane, Fort Collins, Colorado 80526

Chair's Introduction—1:15

Invited Papers

1:20

5pNS1. The effect of rail sources on noise-sensitive construction in Chicago. Brian L. Homans (Shiner + Assoc., Inc., One N. Franklin, Ste. 2025, Chicago, IL 60606, bhomans@shineracoustics.com)

Because of Chicago's dominance as a rail center in the 19th and 20th centuries and the miles of track present, parcels of land available for residential and other noise-sensitive construction frequently lie adjacent to freight, commuter, and Chicago Transit Authority (CTA) tracks. As A-weighted sound levels commonly exceed 105 dBA underneath CTA elevated tracks, special methods of noise abatement for high-rise and other construction to address high noise levels are necessary. Also discussed will be methods of establishing acoustical criteria appropriate for commercial and residential interior spaces.

1:40

5pNS2. 24 hours of noise in a large city, problems and solutions. Leslie Blomberg (Noise Pollution Clearinghouse, P.O. Box 1137, Montpelier, VT 05601)

This paper begins by comparing a contemporary urban noise study to studies done in the 1970s. Several changes in the urban soundscape are noted, as well as areas in which progress has and has not been made in quieting urban soundscapes. Finally, the challenges to quieting urban areas are presented and potential solutions are identified.

2:00

5pNS3. Archaeological acoustics—A guide to trends in community noise levels. Daniel R. Raichel (Eilar & Assoc., 2727 Moore Ln., Fort Collins, CO 80526, draichel@eilarassociates.com), Bennett M. Brooks (Brooks Acoust. Corp., Vernon, CT 06066), and David Lubman (David Lubman & Assoc., Westminster, CA 92683)

Archaeological acoustics may be defined as assessment of the acoustical situations of the past by the use of scientific methods in examining the available remaining evidence. These principles may be applied to the study of historic trends in community noise by measuring the types of noise sources prevalent at the time, estimating the noise level of the environment on the basis of information as to types and number of noise sources, amount of traffic and industry, types of road surfaces, and layout of surrounding buildings and local landscaping. For instance, the clopping of horses and wagons passing by can be measured with modern instruments at current-day rodeos, parades, and exhibitions. The noise levels can be transposed to a bygone scene through acoustical analyses, possibly with computer-aided methodologies. Archives can be consulted for population density, details of buildings that existed at the time, paving types of streets, presence of sidewalks, amount of commercial and pedestrian traffic, etc. Estimates of the noise levels in major cities for specific periods of time can lead to detection of noise trends over a period of centuries. Several noise trend estimate examples are illustrated.

2:20

5pNS4. Evaluation of the noise pollution in urban parks of Curitiba, Brazil. Andressa C. Ferreira, Fabiano B. Diniz, Elaine C. Paz, and Paulo T. Zannin (Environ. Acoust. Lab., Dept. of Mech. Eng., UFPR, Brazil)

This work shows a study about the noise pollution found in six urban parks of Curitiba, Paran, Brazil. The equivalent noise levels (Leq) have been measured in points spread throughout the park, and interviews have been conducted with some park visitors. It has been found out that 52.48% out of the measurement sites did not satisfy the Municipal Law no. 10,625, which states the noise emission level of 55 dB(A) as the limit value for green areas. The results of the questionnaires applied to the local visitors have showed that 39% out of the interviewed people used to visit the park every day and that 75% out of them seek for the realization of a physical activity. During the realization of their activities in the parks, 22% out of the interviewed people pointed to the noise pollution as the source of annoyance and 28% out of them pointed the local security. In this sense, it has been verified that half of the analyzed parks were inserted in acoustically polluted areas, which incurs a real state depreciation in their vicinities.

2:40–2:45 Break

2:45

5pNS5. Public acceptance of urban rotorcraft operations. Michael A. Marcolini, Clemans A. Powell, and Joe W. Posey (NASA Langley Res. Ctr., Hampton, VA 23681)

Even though tiltrotor operations from city center to city center could greatly shorten travel times over moderate ranges, public opposition to intense urban rotorcraft activity has kept this possibility from being realized. One significant factor in this opposition is rotorcraft noise. Over the last 25 years, NASA has explored the subjective response to rotorcraft noise and developed low noise design concepts and noise abatement flight procedures. While low noise designs can be applied for future rotorcraft, this is not an effective near-term means of reducing rotorcraft noise, because of the costs associated with replacement of helicopter rotor blades. Recent noise abatement research, which has been focusing on the development of tools and techniques to facilitate the design of quieter flight procedures for existing vehicles, has much more immediate application. While very little subjective response work has occurred recently, prior work at NASA in this area from the 1970s and 1980s is discussed. Lastly, thoughts on future research areas that might help improve the public acceptance of rotorcraft will be described.

3:00

5pNS6. Study of the daily environmental urban noise levels in San Juan, Puerto Rico. José A. Alicea-Pou, Olga Viñas-Curiel, Wanda Cruz-Vizcarrondo, Daniel Hernandez-Dávila (Noise Control Area, Environ. Quality Board of Puerto Rico), and Jorge Rocaford (School of Architecture, Univ. of Puerto Rico)

The study of the urban environmental noise levels in the city of San Juan is one of the first steps into understanding the noise behavior of the largest and denser populated zones in Puerto Rico. For the year 2002 San Juan had 36% of all citizen noise complaints reported for the entire island. The aim of the study was to monitor for 24 continuous hours the levels of urban sounds and noise for different locations in the city and describe the general tendencies regarding those levels for residential, commercial, recreational or industrial areas. The sampling was done with four (4) units of the Norsonic NOR-121 sound analyzer that were installed temporarily for 24 hours in multiple locations chosen randomly around the city. The study is expected to continue during the next 2 years. The preliminary data showed that the noise levels in residential areas with a high traffic routine fluctuated in levels that were consistently high all day and most of the night, dropping around 1:00 a.m. and rising again around 5:00 a.m. The results of this and other projects would be used to build and implement a National Noise Plan for Puerto Rico for 2005–2007.

3:15

5pNS7. Study on the correlation between the old city and the newly developed city regarding audible sound types and sound levels. Takeshi Tokashiki (Faculty of Eng., Univ. of the Ryukyus, 1 Senbaru Nishihara cho Nakagami gun Okinawa PRef. Japan, tkashiki@tec.u-ryukyu.ac.jp), Yasuhiro Yamashita, and Naoki Takagi (Shinshu Univ., Nagano, Japan)

Sound-level data are the most important among the appraisal data of environmental noise. However, to appraise diversified environment like that of the present times, a more detailed appraisal on environmental noise could be obtained by conducting subjective appraisal in addition to the physical noise level. In this research, we have surveyed subjective appraisal, audible sound types and sound level (LA_{eq}) along boulevard, commercial, and residential areas in the old city and the newly developed city. The subjective appraisal by examinees shows that of the 100% of noise along boulevards, traffic sound occupies 80%; of the 60% of noise in commercial areas the traffic sound occupies 65%; and of the 40% of noise in residential areas the traffic sound occupies 50%. We have studied the correlation between the result of the same surveyed in the old city and the newly developed city. There have been fewer study cases about the correlation between sound types and subjective appraisal; therefore, in this paper we have studied this correlation.

3:30

5pNS8. Comparative analysis of the urban noise between two different areas in the city of Curitiba, PR. Elaine C. Paz, Andressa C. Ferreira, and Paulo T. Zannin (Environ. Acoust. Lab., Dept. of Mech. Eng., UFPR, 81531-990, Curitiba, PR, Brazil)

The purpose of this work is to analyze the urban noise perception comparatively in the inhabitants of a residential area (neighborhood) and a mixed area (center), in the city of Curitiba, PR, in order to characterize two different situations: (1) acoustically ideal urban environment; and (2) acoustically polluted urban environment. For that, subjective and objective evaluations were accomplished, where an aleatory sample of each area was submitted to a survey. In the objective evaluation, the medium equivalent sound levels calculated were 53.50 dB(A) and 72.90 dB(A) for the neighborhood and center, respectively. The parameters used for comparison of the calculated medium equivalent sound levels where the values of 55.00 dB(A) (Municipal Law No. 10.625) and 65.00 dB(A) (WHO), in the period of the day for residential areas. The interpretation of the subjective results verified that the central zone inhabitants have an annoyance perception bigger than the residential zone inhabitants. The interpretation of the objective results classified the neighborhood and center areas as acoustically control zone and acoustically polluted zone, respectively, according to the adopted parameters. Starting from the comparison between these two areas, it was defined that both can be classified as reference factor for other evaluations.

Session 5pPA

Physical Acoustics: Wave Propagation: Inhomogeneous Media

Emmanuel Bossy, Chair

Department of Aerospace and Mechanical Engineering, Boston University, 110 Cummington Street, Boston, Massachusetts 02215

Contributed Papers

1:00

5pPA1. Ultrasonic inhomogeneous waves: Three decades of fascination. Nico F. Declercq (Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, nicof.declercq@ugent.be), Rudy Briers (KATHO, Torhout, Belgium), Joris Degrieck (Ghent Univ., B-9000 Ghent, Belgium), and Oswald Leroy (IRC-KULAK, Krotrijk, Belgium)

Inhomogeneous waves are mathematically described as pure homogeneous plane waves except for the important fact that all wave parameters, i.e., wave vector, polarization vector and frequency, can be complex valued. Even though the existence of a complex wave vector was known for a long time, Henry Cooper was the first to study such waves and their interaction with interfaces. It was not until 1980 that researchers became aware of the fact that inhomogeneous waves are mandatory phenomena in the description of reflection/transmission phenomena between lossy media. Later, it was shown that inhomogeneous waves also formed natural building blocks of bounded beams and formed a physical explanation for the Schoch effect. This paper describes this development from a historical point of view and describes scientific properties (e.g., their polarization, dispersion) whenever needed in order to understand the evolution from 1980 until the present. All topics of inhomogeneous wave research are taken into account, such as waves in viscoelastic solids and liquids, thermo-viscous liquids and solids, anisotropic viscoelastic materials, periodically rough materials, the features of complex frequency and the experimental generation of inhomogeneous waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

1:15

5pPA2. Acoustic pulse propagation through stable atmospheric boundary layer: Theory and experiment. Igor Chunchuzov, Sergey Kulichkov, Alexander Otrezov, and Vitaly Perepelkin (Obukhov Inst. of Atmospheric Phys., 3 Pyzhevskii, Moscow, oksana@achilles.net)

Mesoscale wind speed and temperature fluctuations with periods from 1 min to a few hours significantly affect a variability and turbulent regime of stable atmospheric boundary layer (ABL). Their statistical characteristics are still poorly understood, although the knowledge of such statistics is required when modeling sound propagation through stable ABL. Several field experiments have been conducted to study the influence of mesoscale wind speed fluctuations on acoustic pulse propagation in stable ABL. The results of these experiments are presented in this work. A special acoustic source was used to generate acoustic pulses due to a detonation of air-propane mixture with a repetition period of 1 min or 30 s. The mean wind speed profiles and mesoscale wind fluctuations were measured by Doppler sodar up to a height of 300 m, and by anemometers placed on a 56-mast. From the measurements of the pulse travel time fluctuations at different distances from the source the statistical characteristics of the mesoscale wind fluctuations such as frequency spectra, coherences, horizontal phase speeds and scales have been obtained. Some of the obtained results are interpreted with a recently developed model of internal wave spectrum in a stably stratified atmosphere.

1:30

5pPA3. Time-domain equations for sound propagation in or reflection from rigid porous media. Vladimir E. Ostashev (NOAA/Environ. Technol. Lab., 325 Broadway, Boulder, CO 80305), D. Keith Wilson (U.S. Army Engineer Res. and Development Ctr., Hanover, NH 03755), and Sandra L. Collier (U.S. Army Res. Lab., Adelphi, MD 20783)

In acoustics, many absorbing materials can be considered as porous media with rigid frames. Therefore, studies of sound propagation in or reflection from rigid porous media are important in many applications. These studies are usually done in the frequency domain where equations describing sound propagation in or reflection from porous media are well established. However, there are also problems where such equations are needed in the time domain. Examples include studies of material properties by acoustic impulses and finite-difference time-domain (FDTD) simulation of sound propagation in the presence of absorbing surfaces. In this paper, using Wilson's relaxation model for acoustical properties of porous media [D. K. Wilson, Appl. Acoust. **50**, 171–188 (1997)], we derive a closed set of time-domain, integro-differential equations for the sound pressure and particle velocity in rigid porous media. In the limiting cases of low and high frequencies, these equations coincide with those known in the literature. Furthermore, in these limiting cases, using the relation model, we derive time-domain boundary conditions for sound reflection from a porous medium. This result can be used in FDTD simulation of sound propagation outdoor. [Work supported by the DoD High Performance Computing Modernization Office and ERDC-CRREL.]

1:45

5pPA4. Frequency dependent *P*-wave speed in a porous medium with aligned fractures. Miroslav Brajanovski, Boris Gurevitch (Dept. of Exploration Geophys., Curtin Univ. of Technol., G.P.O. Box U1987, Perth WA 6845, Australia), and Michael Schoenberg (CSIRO Petroleum, Bentley WA 6102, Australia)

In an elastic medium, fractures can be modeled as thin layers, the elastic stiffnesses of which approach zero as the volume fraction of the fractures $h_f \rightarrow 0$. This yields linear slip theory [M. Schoenberg, J. Acoust. Soc. Am. **68**, 1516–1521 (1980)], shown to be a robust way to account for the acoustic effect of fracturing. From Norris' [J. Acoust. Soc. Am. **94**, 359–370 (1993)] dispersion relation for alternating porous layers, fractures must be modeled similarly in porous media. However, another fracture parameter of great interest is permeability. If fracture permeability is taken to be $\mathcal{O}(h_f^{-1})$ or taken to be independent of h_f, one arrives at a dispersion relation for the fast *P*-wave dependent on porous background properties and a real excess compliance which takes the fractures into account. However, if fracture permeability is assumed to be small, and is taken to be $\mathcal{O}(h_f)$, the *P*-wave dispersion dependence on background parameters remains the same, but the term accounting for the fractures is frequency dependent. Only in the zero frequency limit does this result agree with that of the other two cases.

2:00

5pPA5. Acoustical properties of gravel. Keith Attenborough and Olga Umnova (Dept. of Eng., The Univ. of Hull, Cottingham Rd., Hull HU6 7RX, UK)

Gravel is an example of a rigid-porous granular material. It is available with several mean stone sizes and hence with a range of flow resistivity. The flow resistivity of gravel varies significantly with flow velocity. Data for low- and high-amplitude impedance are presented and compared with predictions of linear and nonlinear theories based on the Johnson-Allard model for rigid-porous media. Comparisons are made also between data and predictions for shock wave reflection and transmission at single- and multiple-layer gravel surfaces including recent data obtained with laser-generated acoustic shocks. Tolerable agreement is obtained between data and predictions. It is found that a low flow resistivity gravel layer has a reflection coefficient that has a minimum as the incident pressure is increased. A layered system that offers the lowest reflection coefficient at linear sound pressures does not continue to do so as the incident pressure is increased. [Work supported by USACE ERDC BT25 program.]

2:15

5pPA6. Is a moving average field a proper macro scale measure? John J. McCoy (The Catholic Univ. of America, Washington, DC 20064)

A number of derivations of the Biot equations governing the acoustics of fluid-filled, porous solids, which are based on a more complete formulation that applies to all length scales, accept moving averages of the response fields that enter the more complete formulation as the response fields that enter the Biot equations. This raises a question: Does a moving average field incorporate *only* macro scale variation? A moving average field is presented as one extreme of a class of fields that is formed from sets of discrete local spatial averages. The set of local spatial averages for a moving average field is accomplished for locations that are separated by a vanishing distance. The opposite extreme of a field of local averages is formed from a set of local averages accomplished for locations that are separated by a distance equal to a linear measure of the region of the spatial average. Explicitly demonstrated is that the moving average of a field that contains both macro and micro scale variation *will itself contain* both macro and micro scale variation. The relative suppression of the micro scale variation compared to the macro scale variation, which obtains in one representation of the moving average field, is only apparent; the micro scale variation can be recovered by an appropriate signal processing. This is in contrast to a wavelet-defined field of local averages, an example of the other extreme, for which the suppression of the micro scale variation is absolute. The issue is significant for a derivation of a prediction model that purports to output macro scale response fields.

2:30

5pPA7. Doublet mechanics in acoustics. Christopher N. Layman, Jr. and Junru Wu (Dept. of Phys., Univ. of Vermont, Cook Physical Sci. Bldg., 82 University Pl., Burlington, VT 05405-0125, jwu@zoo.uvm.edu)

Doublet mechanics (DM) presents a micromechanical approach to materials, whose properties allow for a granular characterization. In DM a pair (doublet) of interacting granules represents the fundamental unit of the material; these two granules are either in contact or connected via a thin adhesive layer. The actions of a doublet are described via a convergent Taylor series, which can be truncated at any level dependent upon how discrete the system is. This is the flexibility of DM, meaning it allows for the use of a single framework to describe materials from low discreteness (i.e., animal tissue) to high discreteness (i.e., solids). Some preliminary work [J. Liu and M. Ferrari, Dis. Markers **18**, 175–183 (2002); J. Wu, C. Layman, and J. Liu, J. Acoust. Soc. Am. (in press)] has produced encouraging results in the application of DM to the ultrasonic characterization of tissue.

2:45

5pPA8. Computation of the dynamic thermal properties of a three-dimensional unit cell of porous media by Brownian motion simulation. Camille Perrot (ENTPE and GAUS), Xavier Olny (ENTPE, DGCB URA CNRS 1652, Vaulx-en-Velin Cedex, France), Raymond Panneton, and Richard Bouchard (GAUS, Univ. de Sherbrooke, Canada)

Acoustic dissipation in porous media is mainly due to viscous and thermal mechanisms that occur in the pores of the microstructure. The purpose of this study is the determination of the macroscopic dynamic acoustic bulk modulus and thermal permeability of real foams from a local scale approach. To achieve this goal, two distinct steps are followed. First, the local geometry of a real foam is obtained using computed microtomography (μCT), then a periodic and regularly paving space tetrakaidecahedron cell is identified from the microstructure. Second, the heat equation is solved for the geometrical model. The paper provides a three-dimensional application of the efficient simulation technique of Brownian motion proposed by Torquato *et al.* for steady state diffusion-controlled problems [Appl. Phys. Lett. **55**, 1847–1849 (1989)] and adapted by Lafarge [Poromechanics **II**, 708 (2002)] in a bi-dimensional case. The influence of the model's microstructural details (anisotropy, and struts junction and cross-section) on the macroscopic properties are studied. The predictions of the macroscopic properties using this local scale approach are then compared to experimental measurements.

3:00

5pPA9. Generation of subharmonic and difference frequency acoustic waves in a water-saturated sandy sediment. Byoung-Nam Kim, Kang Il Lee, Suk Wang Yoon, and Bok Kyoung Choi (Dept. of Phys., Sung KyunKwan Univ., Suwon 440-746, Republic of Korea, swyoon@skku.ac.kr)

Generation of subharmonic and difference frequency acoustic waves in a water-saturated sandy sediment was investigated. Subharmonic frequency acoustic wave was observed to be generated due to the nonlinearity of water-saturated sandy sediment when the fundamental frequency acoustic wave exceeded a certain threshold pressure amplitude. Pressure spectrum level of the subharmonic frequency acoustic wave linearly increased as the driving acoustic pressure amplitude increased. The pressure level was 10 dB higher than the background noise level. Generation of the difference frequency acoustic wave was also observed by the collinear acoustic waves with two different fundamental frequencies. The pressure level of the difference frequency acoustic wave was 5 dB higher than the background noise level. It seems very useful to evaluate the nonlinear parameter of water-saturated sandy sediment without disturbing the sediment. Such nonlinear acoustic responses of water-saturated sandy sediment can be utilized for the diagnosis of marine gassy sandy sediment.

3:15–3:30 Break

3:30

5pPA10. Diffusion of high-frequency energy in fluid-saturated porous media. Eric Savin (Structural Dynam. and Coupled Systems Dept., ONERA, 29 Ave. de la Div. Leclerc, 92322 Châtillon cedex, France)

The modern mathematical theory of microlocal analysis shows that the energy associated with the high-frequency solutions of hyperbolic partial differential equations (such as the wave or the Navier equations) satisfy Liouville-type transport equations, or radiative transfer equations for randomly heterogeneous media. For long propagation times the latter can be approached by diffusion equations. Some classical results of the structural acoustics literature about the heat conduction analogy and the statistical energy analysis of structural dynamics at higher frequencies are recovered in this process. The purpose of this communication is to focus on such a diffusive regime for isotropic, fluid-saturated porous media. More specifically, we have derived the diffusion parameters (transport mean-free path and diffusion constant) for such media. Our model considers Biot's equations of poroelasticity, where thermal and viscous effects are modelized by

dynamic tortuosity and compressibility with singular memory kernels. The macroscopic bulk modulus and density of the dry solid phase are assumed to be homogeneous random processes, while tortuosity and porosity remain constant.

3:45

5pPA11. Surface and plate mode stimulation in piezoelectric materials in the framework of inhomogeneous wave theory. Nico F. Declercq, Joris Degrieck (Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, nicof.declercq@ugent.be), and Oswald Leroy (IRC-KULAK, Kortrijk, Belgium)

Inhomogeneous waves are already a well established concept in ultrasonics. Whenever damping is present, such waves become inevitable when reflection and transmission phenomena have to be described. The concept of inhomogeneous waves has been established in the study of acoustics in liquids, isotropic solids and also anisotropic solids. It has been shown before that inhomogeneous waves are much better in stimulating surface waves and plate modes than plane waves and general bounded beams. Nevertheless the concept has not been introduced yet in the field of piezoelectric media. The current paper describes the interaction of inhomogeneous waves with piezoelectric solids of any kind of anisotropy and shows how surface and plate modes can be excited by means of such waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

4:00

5pPA12. Diffraction of plane waves having complex frequency and the excitation of transient leaky Rayleigh waves. Nico F. Declercq, Joris Degrieck (Soete Lab., Ghent Univ., Sint Pietersnieuwstraat 41, B-9000 Ghent, Belgium, nicof.declercq@ugent.be), and Oswald Leroy (IRC-KULAK, Kortrijk, Belgium)

Harmonic plane waves are not advantageous for stimulation of leaky Rayleigh waves whereas bounded beams are better suited. Furthermore, past studies have revealed that it is even better to apply incident harmonic inhomogeneous waves, both on smooth and on rough surfaces. If the right inhomogeneity is used, very strong excitation is to be expected. However, inhomogeneous waves are difficult to generate and are far from practical outside the laboratory boundaries. The current paper describes how incident harmonic plane waves having a complex frequency are also excellent tools for stimulating leaky Rayleigh waves, if the technique is combined with the use of a periodically rough surface. Such transient plane waves are much more practical to generate experimentally whence the usability of the technique is more realistic compared with the classical use of harmonic inhomogeneous waves. [Work supported by The Flemish Institute for the Encouragement of the Scientific and Technological Research in Industry (I.W.T.).]

4:15

5pPA13. Nonlinear modal method of crack localization. Lev Ostrovsky (Zel Technologies/NOAA ETL, Boulder, CO), Alexander Sutin (Stevens Inst. of Technol., Hoboken, NJ), and Andrey Lebedev (Inst. of Appl. Phys., Nizhniy Novgorod, Russia)

A simple scheme for crack localization is discussed that is relevant to nonlinear modal tomography based on the cross-modulation of two signals at different frequencies. The scheme is illustrated by a theoretical model, in which a thin plate or bar with a single crack is excited by a strong low-frequency wave and a high-frequency probing wave (ultrasound). The crack is assumed to be small relative to all wavelengths. Nonlinear scattering from the crack is studied using a general matrix approach as well as simplified models allowing one to find the nonlinear part of crack volume variations under the given stress and then the combinational wave components in the tested material. The nonlinear response strongly depends on the crack position with respect to the peaks or nodes of the corresponding interacting signals which can be used for determination of the crack position. Juxtaposing various resonant modes interacting at the crack it is possible to retrieve both crack location and orientation. Some aspects of inverse problem solutions are also discussed, and preliminary experimental results are presented.

4:30

5pPA14. Omnidirectional elastic bandgap in finite one-dimensional phononic systems. Betsabe Manzanares-Martínez (CIFUS, Univ. of Sonora, Hermosillo, Sonora 83190, Mexico, mbmm@cajeme.cifus.uson.mx), José Sánchez-Dehesa, Andreas Håkansson (Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), Francisco Cervera (Polytechnic Univ. of Valencia, E-46022 Valencia, Spain), and Felipe Ramos-Mendieta (Univ. of Sonora, Sonora 83190, Mexico)

We have demonstrated experimentally the occurrence of omnidirectional elastic bandgaps in layered periodic structures. To date, omnidirectional reflection is a well-known phenomenon for the case of electromagnetic waves in photonic crystals [Y. Fink *et al.*, Science **282**, 1679 (1998)]. Theoretically a similar behavior of the elastic waves in phononic systems is predictable [D. Bria and B. Djafari-Rouhani, Phys. Rev. E **66**, 056609 (2002)]. For the experiment we used finite samples of alternating layers of Pb and epoxy. The binding medium is nylon, which is a material of high elastic velocity, a favorable condition to find the effect. The thicknesses of the layers were chosen in order to have omnidirectional gap at a few hundreds of kHz. We found good agreement between the experimental results and the theoretical treatment of the transmittance through finite samples and the phononic band structure. The effect is shown only for longitudinal impinging waves. [Work supported by CICyT of Spain and CONACyT of Mexico.]

4:45

5pPA15. Criteria for quasi-shear wavefront triplication in a tranversely isotropic material. Michael A. Schoenberg and Thomas M. Daley (Earth Sci. Div., Lawrence Berkeley Lab, 1 Cyclotron Rd., Berkeley, CA 94720)

Quasi-shear (qSV) wavefront triplication, classically known as shear wave bi-refringence, occurs in TI media when anellipticity parameter $E^2 \equiv BC - A^2$ differs significantly from 0 (equality is the elliptic case), where $A \equiv c_{13} + c_{55}$, $B \equiv c_{11} - c_{55}$, $C \equiv c_3 - c_{55}$ [Helbig and Schoenberg, J. Acoust. Soc. Am. **81**, 1235–1245 (1987)]. The region of the slowness curve corresponding to the triplicating region of the wavefront curve must be concave. Most common hexagonal crystals and all TI media long wavelength equivalent to isotropic layering exhibit positive anellipticity. In general, the exact condition for positive anellipticity triplication (the triplicating region then is between the 3-axis and its normal) requires the solution of a cubic equation [Peyton, *Elastic Wave Propagation in Transversely Isotropic Media* (Martinus Nijhoff, 1983)]; however, a good approximation for triplication is that $E^2 > KBCc_{55}/(c_{11} + c_{33})$, where K is almost always within a few percent of 1.39. For negative anellipticity media, a simpler case, triplication centers about the 3-direction when $A^2 > c_{11}C$ and/or about the normal to the 3-axis when $A^2 > c_{33}B$.

5:00

5pPA16. Generalized paraxial ray tracing derived from Riemannian geometry. David R. Bergman (Dept. of Phys., Saint Peter's College, 2641 Kennedy Blvd., Jersey City, NJ 07306, dbergman@spc.edu)

In 1973 R. White demonstrated that acoustic rays in a generic environment could be identified with the null geodesics of a pseudo-Riemannian manifold. A general set of paraxial (dynamic) ray equations, suitable for three dimensional ray tracing in a generic environment, is derived from the geodesic deviation equation used in general relativity and the geometric transmission loss of a ray bundle modeled from this equation. (When fluid motion is removed the paraxial ray procedure used in seismology emerges from the formalism.) The results are applied to time independent layered media where it is found that the standard ray integrals used in underwater acoustics as well as a generalized version of Snell's

law, originally derived by Kornhauser, emerge naturally as conserved quantities related to symmetries of the metric. Finally, when the results are applied to torsion free rays the deviation equation reduces to a scalar equation and the sectional curvature reduces to a simple expression depending on the derivatives of the sound speed and fluid velocity as well as the ray parameters allowing one to determine ray divergence or the development of focal points simply by checking the sign of a single term in the equation.

FRIDAY AFTERNOON, 28 MAY 2004 IMPERIAL BALLROOM B, 1:00 TO 5:00 P.M.

Session 5pSC

Speech Communication: Poster Session V

Nassima Abdelli-Beruh, Chair
Department of Speech-Language Pathology and Audiology, New York University, 719 Broadway, New York, New York 10003

Contributed Papers

All posters will be on display from 1:00 to 5:00 p.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 1:00 p.m. to 3:00 p.m. and contributors of even-numbered papers will be at their posters from 3:00 p.m. to 5:00 p.m.

5pSC1. Nonsensory factors in speech perception. Rachael F. Holt and Arlene E. Carney (Dept. of Commun. Disord., Univ. of Minnesota, 115 Shevlin Hall, 164 Pillsbury Dr. SE, Minneapolis, MN 55455)

The nature of developmental differences was examined in a speech discrimination task, the change/no-change procedure, in which a varying number of speech stimuli are presented during a trial. Standard stimuli are followed by comparison stimuli that are identical to or acoustically different from the standard. Fourteen adults and 30 4- and 5-year-old children were tested with three speech contrast pairs at a variety of signal-to-noise ratios using various numbers of standard and comparison stimulus presentations. Adult speech discrimination performance followed the predictions of the multiple looks hypothesis [N. F. Viemeister and G. H. Wakefield, J. Acoust. Soc. Am. **90**, 858–865 (1991)]; there was an increase in d by a factor of 1.4 for a doubling in the number of standard and comparison stimulus presentations near d values of 1.0. For children, increasing the number of standard stimuli improved discrimination performance, whereas increasing the number of comparisons did not. The multiple looks hypothesis did not explain the children's data. They are explained more parsimoniously by the developmental weighting shift [Nittrouer *et al.*, J. Acoust. Soc. Am. **101**, 2253–2266 (1993)], which proposes that children attend to different aspects of speech stimuli from adults. [Work supported by NIDCD and ASHF.]

5pSC2. Acoustic correlates of the question–statement contrast in children. Rupal Patel, Mariam Syeda (Dept. of Speech Lang. Pathol. and Audiol., Northeastern Univ., 360 Huntington Ave., Boston, MA 02115, r.patel@neu.edu), and Maria Grigos (New York Univ., New York, NY 10003)

The acoustics of prosodic control in children was studied in 4-, 7-, and 11-year olds using the question–statement contrast. Each child produced the utterances "Show Bob a bot" (voiced consonants) and "Show Pop a pot" (voiceless consonants) ten times each as a question and ten times each as a statement. A total of 40 utterances were analyzed per child. The following acoustic measures were obtained for each word within each utterance: average fundamental frequency ($f0$), peak $f0$, slope of $f0$, average intensity, peak intensity and duration. Preliminary results indicate no significant difference between questions and statements for 4-year olds in both the voiced and voiceless consonant conditions. In contrast, 7- and 11-year olds differentiated questions from statements by increasing average, peak, and slope of f0 as well as increasing the duration of the final syllable. Changes in syllable duration between questions and statements were more pronounced for the utterance with voiced consonants. These findings suggest that the acoustics of prosodic contrasts begins to differentiate somewhere between ages 4 and 7 and is influenced by developmental changes in physiological control and flexibility which may also affect segmental features.

5pSC3. Preference patterns in infant vowel perception. Monika T. Molnar and Linda Polka (1266 Pine Ave. W., Montreal, QC H3G 1A8, Canada, linda.polka@mcgill.ca)

Infants show directional asymmetries in vowel discrimination tasks that reveal an underlying perceptual bias favoring more peripheral vowels. Polka and Bohn (2003) propose that this bias is language independent and plays an important role in the development of vowel perception. In the present study we measured infant listening preferences for vowels to assess whether a perceptual bias favoring peripheral vowels can be measured more directly. Monolingual (French and English) and bilingual infants completed a listening preference task using multiple natural tokens of German /dut/ and /dyt/ produced by a male talker. In previous work, discrimination of this vowel pair by German-learning and by English-learning infants revealed a robust directional asymmetry in which /u/ acts as a perceptual anchor; specifically, infants had difficulty detecting a change from /u/ to /y/, whereas a change from /y/ to /u/ was readily detected. Preliminary results from preference tests with these stimuli show that most infants between 3 and 5 months of age also listen longer to /u/ than to /y/. Preference data obtained from older infants and with other vowel pairs will also be reported to further test the claim that peripheral vowels have a privileged perceptual status in infant perception.

5pSC4. Acoustical study of the development of stop consonants in children. Annika K. Imbrie (MIT, 77 Massachusetts Ave., Rm. 36-545, Cambridge, MA 02139, imbrie@mit.edu)

This study focuses on the acoustic patterns of stop consonants and adjacent vowels as they develop in young children (ages 2.6–3.3) over a 6-month period. The acoustic properties that are being measured for stop consonants include spectra of bursts, frication noise and aspiration noise, and formant movements. Additionally, acoustic landmarks are labeled for

measurements of durations of events determined by these landmarks. These acoustic measurements are being interpreted in terms of the supraglottal, laryngeal, and respiratory actions that give rise to them. Preliminary data show that some details of the child's gestures are still far from achieving the adult pattern. The burst of frication noise at the release tends to be shorter than adult values, and often consists of multiple bursts, possibly due to greater compliance of the active articulator. From the burst spectrum, the place of articulation appears to be normal. Finally, coordination of closure of the glottis and release of the primary articulator is still quite variable, as is apparent from a large standard deviation in VOT. Analysis of longitudinal data on young children will result in better models of the development of motor speech production. [Work supported by NIH Grants DC00038 and DC00075.]

5pSC5. The time course of laryngeal coarticulation in children: First results. Laura L. Koenig (Long Island Univ., Brooklyn, NY and Haskins Labs., 270 Crown St., New Haven, CT 06511)

Previous work has suggested that the degree or extent of coarticulation is more extreme in young children than adults. Research in this area has focused primarily on supralaryngeal aspects of coarticulation, using spectral measures such as formant frequencies and fricative centroids. At the same time, studies of adults have found that laryngeal adjustments for voiceless consonants extend well into neighboring vowels, yielding higher values of open quotient and DC flow, and more symmetical pulse shapes, in vowels flanking voiceless as compared to voiced consonants. The current work investigates laryngeal coarticulation in normally developing, English-speaking 4- and 5-year olds. Inverse filtering of the oral airflow is used to approximate the glottal source signal in utterances containing /VpV, VhV, VbV/ sequences. The /VbV/ utterance, which does not require vocal-fold abduction, serves as a control condition to the voiceless consonants. Voice source (open quotient, speed quotient) and aerodynamic (AC and DC flow) quantities are measured over time, and compared between the children and adult females. Along with adding to our understanding of developmental changes in coarticulation, these data will contribute to the literature on differences in voice source properties between children and adults. [Work supported by NIH.]

5pSC6. Fundamental frequency and intensity variability in young children's speech produced at comfortable effort level. William S. Brown, Jr. and Rahul Shrivastav (Dept. of Commun. Sci. and Disord., Univ. of Florida, Dauer Hall, P.O. Box 117420, Gainesville, FL 32611)

Experiments to study speech production and clinical evaluation of speech often require that participants phonate or produce speech at a comfortable level of effort. It is necessary to know the degree of variability in fundamental frequency ($F0$) and intensity (SPL) across multiple test-sessions when speech is elicited in such a manner. If the intersession variability in these measures is too large, the experimental task may need to be appropriately modified. Although the variability in $F0$ and SPL produced at comfortable effort level across multiple sessions has been reported for normal adult speakers, no such data is available for young children. A vowel, a sentence, and four words were produced at a comfortable level by 15 males and 14 females between the ages of 3 and 4 years, on three separate days, one week apart. The $F0$ and SPL for these tasks were compared across the three test sessions. Results show that there were no significant differences in $F0$ and SPL and suggest that normal-speaking children between the ages of 3 and 4 years show the same degree of variability in their $F0$ and SPL as adults when asked to set their own level of comfort.

5pSC7. A developmental acoustic characterization of English diphthongs. Sungbok Lee (Depts. of Elec. Eng. and Linguist., USC, 3601 Watt Way, GFS 301, Los Angeles, CA 90089-1693, sungbokl@usc.edu), Shrikanth Narayanan, and Dani Byrd (USC, Los Angeles, CA 90089)

Acoustic properties of diphthongs in American English in a /bV/ context were investigated as a function of the age of child speakers (5–18 years, 450 subjects). Duration and the average amount of spectral change per frame were computed as a measure of the contrast between initial and final steady states, and the center position of the spectral transition (normalized for vowel length) was estimated as a measure of timing between the two steady-state portions. Results indicate that duration and temporal variability are significantly larger for children below age 10. The spectral change per frame was larger for younger children, suggesting more spectral contrast between diphthong steady states. Interestingly, the normalized center position of the glide occurs at a position relatively closer to the final portion of diphthongs for /ay/ for younger children; an opposite tendency was observed in the case of /ow/, implying that younger children exhibit delayed or premature spectral transitions. Together these observations suggest the possibility that 5–8-year olds may not produce diphthongs as a unit but, rather, as a concatenation of two vowel targets in which one component is more prolonged than the other and abrupt transitions occur. [Work supported by SBC/TRI, NSF, and NIH.]

5pSC8. Spectral moments versus Bark cepstrum classification of children's voiceless stops. James Polikoff, Jenna Hammond, Jane McNicholas, and H. Timothy Bunnell (Speech Res. Lab, duPont Hospital for Children, 1600 Rockland Rd., Wilmington, DE 19807)

Spectral moments have been shown to be effective in deriving acoustic features for classifying voiceless stop release bursts [K. Forrest, G. Weismer, P. Milenkovic, and R. N. Dougall, J. Acoust. Soc. Am. **84**, 115–123 (1988)]. In this study, we compared the classification of stops /p/, /t/, and /k/ based on spectral moments with classification based on an equal number of Bark cepstrum coefficients. The speech tokens were 446 instances each of utterance-initial /p/, /t/, and /k/ sampled from utterances produced by 208 children 6 to 8 years old. Linear discriminant analysis (LDA) was used to classify the three stops based on four analysis frames from the initial 40 ms of each token. The best classification based on spectral moments used all four spectral moment features and all four time intervals and yielded 75.6% correct classification. The best classification based on Bark cepstrum yielded 83.4% correct also using four coefficients and four time frames. Differences between these results and previous classification results using spectral moments will be discussed. Implications for future research on the acoustic characteristics of children's speech will be considered.

5pSC9. Acoustic characterization of developmental speech disorders. H. Timothy Bunnell, James Polikoff, Jane McNicholas, Rhonda Walter (Speech Res. Lab, duPont Hospital for Children, 1600 Rockland Rd., Wilmington, DE 19807, bunnell@asel.udel.edu), and Matthew Winn (Univ. of Delaware)

A novel approach to classifying children with developmental speech delays (DSD) involving /r/ was developed. The approach first derives an acoustic classification of /r/ tokens based on their forced Viterbi alignment to a five-state hidden Markov model (HMM) of normally articulated /r/. Children with DSD are then classified in terms of the proportion of their /r/ productions that fall into each broad acoustic class. This approach was evaluated using 953 examples of /r/ as produced by 18 DSD children and an approximately equal number of /r/ tokens produced by a much larger number of normally articulating children. The acoustic classification identified three broad categories of /r/ that differed substantially in how they aligned to the normal speech /r/ HMM. Additionally, these categories tended to partition tokens uttered by DSD children from those uttered by normally articulating children. Similarities among the DSD children and average normal child measured in terms of the proportion of their /r/

productions that fell into each of the three broad acoustic categories were used to perform a hierarchical clustering. This clustering revealed groupings of DSD children who tended to approach /r/ production in one of several acoustically distinct manners.

5pSC10. Learning to talk. Piers Messum (Dept. of Phonet., Univ. College, London, UK)

Is imitation a necessary part of learning to talk? The faithful replication by children of such arbitrary phenomena of English as tense and lax vowel properties, ''rhythm,'' and context-dependent VOT's seems to insist that it is. But a nonimitative account of this is also possible. It relies on two principal mechanisms. First, basic speech sounds are learned by emulation: attempting to reproduce the results achieved by other speakers but without copying their actions to do so. The effectiveness of the output provides sufficient feedback to inform the child of the adequacy of its performance and to guide refinement. Second, phonetic phenomena such as those above appear through aerodynamic accommodation. Key elements of this are (a) that speech breathing is a complex motor skill which dominates other articulatory processes during acquisition and starts pulsatile before becoming smooth, and (b) that a child-scale production system imposes constraints on talking which do not operate in the adult speaker. Much of ''the terrible complexity of phonetic patterns'' [J. Pierrehumbert, Lang. Speech **46**, 115–154 (2003)] then becomes epiphenomenal: appearing not as a result of young learners copying phonetic detail that is not linguistically significant, but of them reconciling conflicting production demands while just talking.

5pSC11. Exploring production-perception relationships for 4-year-old children: A study of compensation strategies to a lip-tube perturbation. Lucie Menard (Departement de linguistique et de didactique des langues, UQAM, CP 8888, succ. Ctr.-Ville, Montreal, QC H3C 3P8, Canada, menard.lucie@uqam.ca), Pascal Perrier, and Christophe Savariaux (CNRS-INPG-Universite Stendhal, Grenoble, France)

The relationships between production and perception for 4-year-old children were examined through a study of compensation strategies to a lip-tube perturbation. Acoustic and perceptual analyses of the rounded vowel [u] produced by 12 4-year-old French speakers were conducted under two conditions: in normal condition and with a 15-mm-diam tube inserted between the lips. Recordings of isolated vowels were made in normal condition before any perturbation ($N1$), immediately upon insertion of the tube ($P1$), for each of the next 20 trials in this perturbed condition ($P2$), and in normal condition after the perturbed trials ($N2$). Results of the acoustic analyses reveal speaker-dependent alteration of $F1$, $F2$, and/or $F0$ in the perturbed conditions and after the removal of the tube. For some subjects, the tube introduced very little changes; for some others, an increase of $F2$ was observed in $P1$, which was generally at least partly compensated during the $P2$ repetitions. Perceptual data are used to determine optimal combinations of $F0$, $F1$, and $F2$ (in bark) related to these patterns. The data are compared to a previous study conducted with adults [Savariaux *et al.*, J. Acoust. Soc. Am. **106**, 381–393 (1999)].

5pSC12. Identification and discrimination of Spanish front vowels. Isabel Castellanos and Luis E. Lopez-Bascuas (Universidad Complutense de Madrid, Facultad de Psicologa, Campus de Somosaguas, 28223 Madrid, Spain, lelopezb@psi.ucm.es)

The idea that vowels are perceived less categorically than consonants is widely accepted. Ades [Psychol. Rev. **84**, 524–530 (1977)] tried to explain this fact on the basis of the Durlach and Braida [J. Acoust. Soc. Am. **46**, 372–383 (1969)] theory of intensity resolution. Since vowels seem to cover a broader perceptual range, context-coding noise for vowels should be greater than for consonants leading to a less categorical performance on the vocalic segments. However, relatively recent work by Macmillan *et al.* [J. Acoust. Soc. Am. **84**, 1262–1280 (1988)] has cast doubt on the assumption of different perceptual ranges for vowels and consonants even though context variance is acknowledged to be greater for the former. A possibility is that context variance increases as number of long-term phonemic categories also increases. To test this hypothesis we focused on Spanish as the target language. Spanish has less vowel categories than English and the implication is that Spanish vowels will be more categorically perceived. Identification and discrimination experiments were conducted on a synthetic /i/-/e/ continuum and the obtained functions were studied to assess whether Spanish vowels are more categorically perceived than English vowels. The results are discussed in the context of different theories of speech perception.

5pSC13. Classification of stop consonant place of articulation. Atiwong Suchato (Speech Commun. Group, MIT, Cambridge, MA 02139, atiwong@mit.edu)

In this study we develop an experimental procedure for examining the relative importance of knowledge-based cues for identifying place of articulation for stop consonants. A set of acoustic attributes is selected for place classification of stops: amplitude and energy of burst, formant movement of adjacent vowels, spectrum of noise after the release, and some temporal cues. The ability of each attribute to separate the three places is evaluated by the classification error based on the distributions of its values for the three places, and another quantifier based on F ratio. These two quantifiers generally agree and show how well each individual attribute separates the three places. Linear discriminant function analysis is used to address the relative importance of these attributes when combinations are used. Their discriminating abilities and the ranking of their relative importance to the classification in different vowel and voicing contexts are reported. The overall findings are that attributes relating to the burst spectrum in relation to the vowel contribute most effectively, while formant transition is somewhat less effective. The approach used in this study can be applied to different classes of sound, as well as stops in different noise environments. [Work supported by NIH Grant Number DC 02978.]

5pSC14. Role of experience in eliciting perceptual overshoot with sine waves. Radhika Aravamudhan and John W. Hawks (Kent State Univ., Kent, OH 44242, raravam1@kent.edu)

Previous studies in vowel perception have demonstrated perceptual overshoot (PO) with vowels in CVC contexts. The first experiment investigated the elicitation of PO with speech and nonspeech signals. Boundary estimates were obtained for four continua: (1) steady-state synthetic vowels; (2) synthetic vowel transitions in /wV/ context; (3) sine wave analogs mimicking the steady-state vowels; and (4) sine wave analogs mimicking the transition continuum. Results replicated the findings of previous studies by demonstrating PO with synthetic vowels; however, PO was not elicited with the sine wave analogs. This led to a second experiment, where the role of training on the elicitation of PO with sine wave analogs was investigated. Subjects from the first experiment were divided into two groups. One group was trained to categorize the steady-state sine wave analogs based on their synthetic vowel boundary. Training occurred over 3–6 1-h sessions until subjects achieved 90%-correct identification. The control group received no training. Post-training boundary estimates were obtained for the four original continua from both groups. The results indicated that PO was now elicited with the sine wave analogs for the trained group, but not for the control group. Further implications of these findings will be discussed.

5pSC15. The perceptual magnet effect reflects phonetic context. Sarah Hawkins (Dept. of Linguist., Univ. of Cambridge, Sidgwick Ave., Cambridge CB3 9DA, UK, sh110@cam.ac.uk) and Sarah Barrett Jones (City Univ., London EC1V 0HB, UK)

Two experiments demonstrate that the perceptual magnet effect is context sensitive. In experiment 1a, 24 participants rated goodness of synthetic /u/ in isolation (*ooh*) and in two consonantal contexts, /lu/, /ju/ (*Lou, you*), with nine versions per word, varying in $F2$ frequency. Their most (prototypical) and least (nonprototypical) preferred choices reflected expected differences between words, and individual differences within words. Experiment 1b demonstrated standard perceptual magnet effects for prototype and nonprototype /u/ in the three words. Unlike previous work, each participant discriminated his/her own prototype and nonprototype from experiment 1a, rather than the group mean. Experiment 2a replicated 1a with 40 new participants. Experiment 2b compared discrimination around participants' prototypical $F2$ frequency for /u/ in one word (original) with discrimination around that same frequency in another word (transferred). Different original/transferred sets were heard by four groups (ten participants each): /u/ and /lu/; /lu/ and /u/; /ju/ and /u/; /ju/ and /lu/. Discrimination (d') near the prototype was poorer for original than transferred contexts [for the four comparisons, $t(9)$ ranged between 2.43–3.49, $p<0.025$–0.005]. Thus, the perceptual magnet effect is syllable specific: the vowel prototype for one word need not generalize to another. Implications for perceptual coherence and phonological representation are discussed.

5pSC16. Phonotactic constraints, frequency, and legality in English onset-cluster perception. Elliott Moreton (Dept. of Linguist., Univ. of North Carolina, Dey Hall, Rm. 318, Chapel Hill, NC 27599-3155, moreton@unc.edu)

Phonological context can affect phoneme identification, favoring one response over another [Massaro and Cohen, Percept. Psychophys. **34**, 338–348 (1983)]. It is unclear what is disfavored: infrequency, legality, or phonotactic constraint violation. This study compared bias on an [l]–[w] continuum in two situations. In one, the three factors were deliberately confounded; in the other, frequency and legality were controlled. Stimuli were synthetic syllables ambiguous between [blae bwae dlae dwae] (''bd array'') or [mlae mwae nlae nwae] (''mn array''). In the bd array, [bl dw] are legal, frequent English onsets, while [dl bw] have zero frequency, are illegal, and violate a constraint against same-place onsets; hence, ''d'' responses should facilitate ''w'' responses. The whole mn array is unattested and illegal, but [nl mw] violate the constraint. Simultaneous stop-sonorant judgments were obtained from English listeners for each 6-by-6 array. A mixed-effects logistic-regression model (including a term to model out compensation for coarticulation) was used to measure dependency between stop and sonorant responses. The bd array drew the expected bias against ''dl.'' Bias against ''nl'' in the mn array was considerably smaller, but still present. These results suggest that perceptual bias is jointly determined by constraint violation, and by frequency and/or legality.

5pSC17. Effect of retroflex sounds on the recognition of Hindi stops. Amita Dev, S. S. Agrawal, and D. Roy Choudhary (Ambedkar Inst. of Technol., Shakarpur, Madhuban, Delhi 92, India, amita_dev@hotmail.com)

As development of the speech recognition system entirely depends upon the spoken language used for its development and the very fact that speech technology is highly language dependent and reverse engineering is not possible, there is an utmost need to develop such systems for Indian languages. In this paper we present the implementation of a time-delay neural network system (TDNN) in a modular fashion by exploiting the hidden structure of previously phonetic subcategory network for the recognition of Hindi consonants. For the present study we have selected all the Hindi phonemes for the recognition. A vocabulary of 207 Hindi words was designed for the task-specific environment and used as a database. For the recognition of phonemes a three-layered network was constructed and the network was trained using the backpropagation learning algorithm. Experiments were conducted to categorize the Hindi voiced and unvoiced stops, semivowels, vowels, nasals, and fricatives. A close observation of the confusion matrix of Hindi stops revealed maximum confusion of retroflex stops with their nonretroflex counterparts.

5pSC18. Perception of correlations between acoustic cues in category tuning and speaker adaptation. Lori Holt and Travis Wade (Dept of Psych., Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213, lholt@andrew.cmu.edu)

In English and many other languages, fundamental frequency ($f0$) varies with voicing such that voiced consonants are produced with lower $f0$'s than their voiceless counterparts. This regularity robustly influences perception, such that sounds synthesized or spoken with a low $f0$ are more often perceived as voiced than are sounds with a higher $f0$. This series of studies exploited these observations to investigate category tuning as a function of incidental exposure to correlations among speech cues and adaptation to speaker idiosyncrasies or accent. Manipulation of $f0$ across sets of natural speech utterances produced stimulus sets varying in their inherent $f0$/voicing relationship. Listeners were exposed to these different $f0$/voicing patterns via spoken word and nonword items in a lexical decision task, and their resulting categorization of ambiguous consonants varying in $f0$ and voice onset time (VOT) was measured. The results suggest listeners adapt quickly to speaker-specific cues but also remain influenced by more global, naturally occurring covariance patterns of $f0$ and voicing in English. This pattern contrasts somewhat with studies where idiosyncrasy is represented instead by manipulation of primary, first-order cues to speech sounds, in which listeners are seen to adapt more straightforwardly to the cues they are presented.

5pSC19. On the categorical nature of Korean /pk/ place assimilation. Minjung Son (Linguist. Dept., Yale Univ. and Haskins Labs., 370 Temple St., New Haven, CT 06520-8366) and Alexei Kochetov (Simon Fraser Univ., Burnaby, BC, Canada)

Korean exhibits regressive place assimilation in /pk/ clusters, which has been described as gradient and rate dependent. However, this assumption has empirically only been tested on the basis of air pressure data [Jun, 1996] which does not provide a direct record of articulator movement. The present study examines articulator movement using EMMA. For three Seoul-dialect speakers, stimuli containing /pk/ clusters were elicited word-medially (for words and nonwords) and in a phrase-boundary condition; two rates were employed. Results show that the labial can indeed reduce word medially, rendering [kk]. However, contrary to previous claims, the data demonstrate that reduction in /pk/ is always categorical, although it is optional or stochastic in its occurrence. Substantial interspeaker variation is observed, with the frequency of reduction being higher at fast rate and ranging overall from 6 at both rates and is never gradient. The lack of reduction in nonsense words and in the phrase boundary condition shows that the process is sensitive to lexical properties. The observed tendency for more gestural overlap word medially compared to the phrase-boundary condition supports the hypothesis that gestural overlap plays a role in the origins of place assimilation. [Work supported by NIH.]

5pSC20. Information conveyed by vowels about other vowels. Hector R. Javkin, Elaine Drom, Carol Christie, and Gaston R. Cangiano (Dept. of Linguist. and Lang. Development, San Jose State Univ.)

Rapid adaptation to different speakers has become an important issue in speech recognition (ASR) but has been known in human listeners for a long time. Ladefoged and Broadbent [J. Acoust. Soc. Am. **29**, 98–104 (1957)] demonstrated that human perception of synthesized vowels occurring in monosyllabic words (bit, bet, bat, but) can be changed by changing the formants of an introductory phrase (Please say what this word is).

These stimuli meant that the introductory phrase ranged over the same portion of the vowel space (front vowels, or relatively high $F2$), thus facilitating listeners adaptation. To further test the limits of human adaptation, we replicated the experiment keeping the same words, substituting introductory phrases consisting of back (low $F2$) vowels, and maintaining a similar level of low-quality synthesized speech. The effects are difficult to replicate with natural or high-quality synthetic speech. However, we will suggest that low quality speech is analogous to the low-dimensionality representation of speech of many ASR front ends, which discard, for example, information as to the higher formants. Therefore, these findings are relevant to the use of adaptation in improving ASR. [Work supported by a Faculty Small Grant from San Jose State University.]

5pSC21. On the relation of apparent naturalness to phonetic perceptual identification of vowels. Stephanie Wissig, Daria F. Ferro, Kate Liberman, Jill Thompson, and Robert E. Remez (Dept. of Psych., Barnard College, 3009 Broadway, New York, NY 10027)

A set of synthetic test syllables was created varying the attributes of naturalness and phonetic vowel height and advancement. These acoustic items were used in tests calibrating the relation of naturalness to phonetic perceptual resolution. Two acoustic methods were used to create naturalness variants: (1) variation in the excitation of the synthetic voicing source and (2) variation in the bandwidth of the formant centers. A naturalness tournament was composed of items drawn from the test series, and the sensitivity of perceivers to the vowel contrast was estimated with the cumulative d' across the series in identification tests. These outcomes reveal different effects of each acoustic manipulation: an independence of the outcome of the naturalness tournament and the measures of phonetic sensitivity in one case and a close relation between naturalness and phonetic resolution in the other. Together, the findings show that intelligibility and naturalness can be either orthogonal or contingent aspects of speech perception. These measures offer a tool to understand rule-based and exemplar-based components of phonetic perception. [Research supported by NIH (DC00308).]

5pSC22. Spectral frequency modulation in vowel identification. Chang Liu and David A. Eddins (Psychoacoustic Lab., Dept. of Commun. Disord. and Sci., Univ. at Buffalo, Buffalo, NY 14214)

Psychophysical and physiological studies have demonstrated selectivity for spectral envelope frequency (also termed spatial frequency) in the auditory system, suggesting that auditory perception of complex sounds might be based on spectral envelope channels. The present study investigated relative contribution of different spatial frequencies to vowel identification. Twelve naturally-spoken American-English vowels were presented at 70 dB SPL. In different conditions, vowel stimuli were subjected to various degrees of low-pass and high-pass filtering in the spatial frequency domain, in effect, altering their spectra. Identification performance for the vowels with and without spatial frequency filtering was estimated for normal-hearing listeners. Results indicated that vowel identification performance was progressively degraded as spatial-frequency components were removed. Results will be interpreted in terms of spatial frequency regions most important to specific vowel categories. The specificity and universality of spatial frequency modulations in vowel identification across different vowel categories will be discussed.

5pSC23. Gradient sensitivity to acoustic detail and temporal integration of phonetic cues. Bob McMurray, Meghan A. Clayards, Richard N. Aslin, and Michael K. Tanenhaus (Univ. of Rochester, Rochester, NY 14627)

Speech contains systematic covariation at the subphonemic level that could be used to integrate information over time (McMurray *et al.*, 2003; Gow, 2001). Previous research has established sensitivity to this variation: activation for lexical competitors is sensitive to within-category variation in voice-onset-time (McMurray *et al.*, 2002). This study extends this investigation to other subphonemic speech cues by examining formant transitions (r/l and d/g), formant slope (b/w) and VOT (b/p) in an eye-tracking paradigm similar to McMurray *et al.* (2002). Vowel length was also varied to examine temporal organization (e.g., VOT *precedes* the vowel). Subjects heard a token from each continua and selected the target from a screen containing pictures of the target, competitor and unrelated items. Fixations to the competitor increased with distance from the boundary along each of the speech continua. Unlike prior work, there was also an effect on fixations to the *target*. There was no effect of vowel length on the d/g or r/l continua, but rate dependent continua (b/w and b/p) showed length effects. Importantly, the temporal order of cues was reflected in the pattern of looks to competitors, providing an important window into the processes by which acoustic detail is temporally integrated.

5pSC24. A novel acoustic way of demarcating diphthong elements in Thai. Rungpat Roengpitya (Dept. of Linguist., Faculty of Arts, Chulalongkorn Univ., Bangkok 10330, Thailand, rungpatr@hotmail.com)

Thai has three phonemic (falling) diphthongs /ia/, /i‘a/, and /ua/ and a set of the so-called rising diphthongs in Thai /ai/ /aj/ and /au/ /aw/. In previous (nonacoustic) literature, falling diphthongs are short when followed by a glottal stop. However, recent acoustic studies revealed that falling diphthongs in Thai are short in closed syllables, but long in open syllables. In the same literature, each diphthong was measured for the durations of its first vocalic element, the transition, and the second vocalic element, without demarcating the offset of the first vocalic element and the onset of the second vocalic element. This paper has the major aim of applying a new technique for making a plausible way to mark where the first vocalic element ends and the second vocalic element starts, not only for the set of falling diphthongs but also for the so-called rising diphthongs. Moreover, this new technique would help to find out whether the rising diphthongs in Thai have the combinations of two-vowel qualities or a vowel and a glide. It is hoped that this new method will be useful for the future acoustic studies in other languages of the world. [Work supported by new-lecturer grants, Chulalongkorn University.]

5pSC25. Obtaining a palatal trace for ultrasound images. Melissa A. Epstein, Maureen Stone, Marianne Pouplier (Vocal Tract Visualization Lab, Biomed. Sci., Univ. of Maryland Dental School, Rm. 5A12, 666 W. Baltimore St., Baltimore, MD 21201, mae001@dental.umaryland.edu), and Vijay Parthasarathy (Johns Hopkins Univ., Baltimore, MD 21218)

This paper presents methods for collection and display of the palate with ultrasound, for use as a reference for tongue movements. Ultrasound does not usually capture structures other than the tongue, because the air above the tongue in the vocal tract reflects the ultrasound beam back to the transducer. However, when the tongue touches the palate, the ultrasound beam is transmitted through the soft tissue until it reaches and is reflected by the palatine bone. The tongue touches the palate during swallowing and some speech sounds. The palate contour can be traced from these images. The paper presents a corpus of speech and swallowing tasks that can be used to create a full palatal trace. The corpus is tested on a subject for whom it is easy to collect palatal images and a subject for whom it is difficult to collect palatal images. The availability of a palate will enhance our ability for data quantification from ultrasound images. In combination with tongue contours, the palate contour allows the computation of lin-

guistically important measures, such as the constriction degree, area functions, and $L2$ norms. [Work supported by NIH RO1-DC01758 and T32-DE07309.]

5pSC26. Effects of human fatigue on speech signals. Catherine Stamoulis (Cambridge Sci. Solutions, 140 Marlborough St., Boston, MA 02116)

Cognitive performance may be significantly affected by fatigue. In the case of critical personnel, such as pilots, monitoring human fatigue is essential to ensure safety and success of a given operation. One of the modalities that may be used for this purpose is speech, which is sensitive to respiratory changes and increased muscle tension of vocal cords, induced by fatigue. Age, gender, vocal tract length, physical and emotional state may significantly alter speech intensity, duration, rhythm, and spectral characteristics. In addition to changes in speech rhythm, fatigue may also affect the quality of speech, such as articulation. In a noisy environment, detecting fatigue-related changes in speech signals, particularly subtle changes at the onset of fatigue, may be difficult. Therefore, in a performance-monitoring system, speech parameters which are significantly affected by fatigue need to be identified and extracted from input signals. For this purpose, a series of experiments was performed under slowly varying cognitive load conditions and at different times of the day. The results of the data analysis are presented here.

5pSC27. HOCUS: The Haskins optically-corrected ultrasound system for measuring speech articulation. D. H. Whalen, Khalil Iskarous, Mark K. Tiede, and David J. Ostry (Haskins Labs., 270 Crown St., New Haven, CT 06511)

The tongue is the most important supralaryngeal articulator for speech, yet, because it is typically out of view, its movements have been difficult to quantify. Here is described a new combination of techniques involving ultrasound in conjunction with an optoelectric motion measurement system (Optotrak). Combining these, the movements of the tongue are imaged and simultaneously corrected for motion of the head and of the ultrasound transceiver. Optotrak's infrared-emitting diodes are placed on the transceiver and the speakers head in order to localize the ultrasound image of the tongue relative to the hard palate. The palate can be imaged with ultrasound by having the ultrasound signal penetrate a water bolus held against the palate by the tongue. This trace is coregistered with the head and potentially with the same talker's sagittal MR image, to provide additional information on the unimaged remainder of the tract. The tongue surface, from the larynx to near the tip, can then be localized in relationship to the hard palate. The result is a fairly complete view of the tongue within the vocal tract at sampling rates appropriate for running speech. A comparison with other imaging vocal tract systems will be presented. [Work supported by NIH Grant DC-02717.]

5pSC28. Movie MRI at five frames a second for evaluation of speech and swallowing. Masanobu Kumada (Dept. of Otorhinolaryngol., Hospital, Natl. Rehabilitation Ctr. for the Disabled (NRCD), 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan), Koichi Mori, Yasoichi Nakajima (NRCD, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan), and Seiji Nozaki (Toshiba Medical Systems Corp.)

Magnetic resonance imaging (MRI) is a noninvasive imaging method that is widely used in the medical field. One of the limitations of MRI is its low time-resolution; images of MRI are usually obtained as still images. Here we introduced a newly developed method of "movie" MRI with high time-resolution at five images a second. Its good application would include study and evaluation of speech and swallowing. Instrument: MRT-2001 XG with Software ver.5.5. (Toshiba). Coil: head QD coil for the tongue; CTL ARRY coil (3ch) for the neck, Imaging condition: sagital FFE2D; TR=2.8 ms; TE=1.2 ms; FA=10 deg; Matrix=64×128; ST=10 mm; NAQ=1 AV; FOV=23×35; ROAFI; sequential acquisition. Maximum imaging length=51 s. A healthy Japanese male (Tokyo dialect speaker, 39 years). Task: Repetitive utterance of /tenten . . . / and intentional swallowing of saliva. Results: In the task of /tenten . . . ./, we could detect, in temporal order, velopharyngeal (VP) closure, opening of the tongue-palate (TP) closure, VP opening, TP closure, VP closure, and so on. In the swallowing task of saliva, we could detect movement of the tongue for conveying saliva to the pharynx, VP closure, backward movement of the tongue root, elevation and descent of the larynx in this temporal order. Our "Movie MRI" seemed promising for the noninvasive evaluation of speech and deglutition. Appropriate materials for swallowing evaluation will be presented.

5pSC29. Anterior tongue and jaw movement in sVd words. Richard S. McGowan (CReSS LLC, 1 Seaborn Pl., Lexington, MA 02420, cressllc@earthlink.net)

The relations between jaw and tongue movements were examined using words of the form sVd for eight different speakers from the X-Ray Microbeam Speech Production Database. Measurements were examined at the maximum speeds during the release and during the closure of the tongue blade and tongue body. For nonhigh vowels the tongue blade traveled in the same direction as the jaw during release, and, to a lesser extent, the same was true during closure. Further, the magnitude of the projection of the tongue blade velocity onto the direction of the jaw movement was often large compared with the speed of the jaw. There was less consistency in the relation between tongue body and jaw movement. These results indicate that the jaw and tongue movements are not rigidly coupled. Rather the jaw, which can provide a hard boundary for the tongue, is getting out of the tongue's way during release and following the tongue on closure for subsequent bracing. [Work supported by Grant NIDCD-01247 to CReSS LLC.]

5pSC30. Tongue–jaw kinematic correlates of /s/ spectra. James S. Dembowski and Richard K. Crumb (Commun. Disord., State Univ. of New York at New Paltz, HUM 14a, 75 S. Manheim Blvd., New Paltz, NY 12561-2499, dembowsj@newpaltz.edu)

Frequencies of spectral peaks for fricatives are determined by the size of the resonating cavity anterior to the place of articulatory constriction in the upper vocal tract. For /s/, this cavity size may be altered through anterior–posterior (a–p) movements of the tongue blade forming the constriction, changes of jaw height, and degree of lip protrusion. With respect to intensity, modeling studies suggest that intensity of fricative spectral peaks may be related to degree of articulatory constriction. These spectral–kinematic relationships have been little studied in natural speech. This study used data from the University of Wisconsin X-Ray Microbeam Database to examine the relationship between spectral peaks and movements of the tongue and jaw in the /s/ productions of one normal speaker. Results showed no relationship between a–p tongue position and frequency of spectral peaks. However, a significant inverse correlation related peak between frequency and jaw opening. Thus, for this speaker jaw height appeared a more important determinant of spectral variability for /s/ than tongue position. Additional results showed a significant relationship between peak intensity and distance of the tongue blade from the palate. These natural speech data will be discussed with respect to models and theories of fricative production.

5pSC31. Correlation between angle of incidence and sliding patterns of the tongue along the palate in Korean velar stops. Jana Brunner (Zentrum fuer Allgemeine Sprachwissenschaft [ZAS] Berlin, Germany, brunner@zas.gwz-berlin.de), Susanne Fuchs (ZAS, Berlin, Germany), Pascal Perrier (Institut de la Commun. Parle, & Univ. Stendhal, Grenoble, France), and Hyeon_Zoo Kim (Dankook Univ., Seoul, Korea)

In former studies, it has been hypothesized that the articulatory production of oral stops could result from the interaction between the tongue moving towards a virtual target located above the palate, and the palate. Velar stops, where the tongue slides along the palate during the occlusion

phase, offer a nice experimental framework for further experimental assessments. Indeed, in the framework of the ''virtual target'' hypothesis, the sliding movement should be seen as the continuation of the movement before the occlusion, but constrained by the palate. Hence, relations should exist between the movement characteristics before contact and during the occlusion phase. To test this hypothesis three Korean speakers were recorded via EMA producing /aCV/ sequences with C=/g/, /k'/ and /kh/, V=/a/, /i/ or /u/. The angle between tongue trajectory just before the impact and palatal contour was estimated, and the amplitude of the sliding movement was measured. Preliminary results for two speakers show that these two variables correlate: The greater the angle, the larger the sliding movement. These findings are interpreted as supporting the ''virtual target'' hypothesis. This interpretation will be verified by simulations using a 2D biomechanical tongue model [Payan and Perrier, Speech Commun. **22** (1997)].

5pSC32. Modeling the acoustics of American English /r/ using configurable articulatory synthesis (CASY). Heike Lehnert-LeHouillier, Khalil Iskarous, and Douglas H. Whalen (Haskins Labs., 270 Crown St., New Haven, CT 06511)

The claim that articulatory variation in /r/ production exhibits systematic tradeoffs to achieve a stable acoustic signal (Guenther *et al.*, 1999) was tested using configurable articulatory synthesis (CASY) and ultrasound data. In particular, the hypothesis was tested that multiple constrictions during /r/ production are necessary to achieve a low enough $F3$. Ultrasound and Optotrak data from four speakers pronouncing /r/ in different vocalic contexts were used to determine where in the vocal tract the tongue gestures are placed. This data was then modeled using CASY parameters and was used to determine how the three gestures in /r/ (labial, palatal, and pharyngeal) contribute to the $F3$ value observed in the speech signal simultaneously recorded with the ultrasound. This was done by varying the degree and location of the lingual constrictions and the degree of the labial constriction and determining the effect on $F3$. It was determined that the three gestures in /r/ contribute in differing amounts to the overall $F3$ lowering. Furthermore, it does not seem that all three gestures are necessary for $F3$ lowering. This lends support to the hypothesis that the goal in /r/ production is the simultaneous achievement of three gestures. [Work supported by NIH Grant DC-02717.]

5pSC33. Formants and musical harmonics matching in Brazilian lied. Beatriz Raposo de Medeiros (Universidade de So Paulo Av Prof. Luciano Gualberto, 403, CEP 01060-970, Sao Paulo, Brazil)

This paper reports a comparison of the formant patterns of speech and singing. Measurements of the first three formants were made on the stable portion of the vowels. The main finding of the study is an acoustic effect that can be described as the matching of the vowel formants to the harmonics of the sung note (A flat, 420 Hz). For example, for the vowel [a], $F1$ generally matched with the second harmonic (840 Hz) and $F2$ with the third harmonic. This finding is complementary to that of Sundberg (1977) according to which the higher the fundamental frequency of the musical note, e.g., 700 Hz, the more the mandible is lowered causing the elevation of the first formant of the sung vowel. As Sundberg himself named this phenomenon, there is a matching between the first formant and the phonation frequency, causing an increase in the sound energy. The present study establishes that the matching affects not only $F1$ but also $F2$ and $F3$. This finding will be discussed in connection with other manoeuvres (e.g., tongue movements) used by singers.

5pSC34. Vowel length in Farsi. Shabnam Shademan (Linguist. Dept., UCLA, 3125 Campbell Hall, Los Angeles, CA 90095)

This study tests whether Farsi vowels are contrastive with respective to length. Farsi has a six-vowel system with three lax vowels and three tense vowels. Both traditional grammarians and modern linguists believe that Farsi tense vowels are longer than lax vowels, and that there are no vowel pairs that contrast only in length. However, it has been suggested that Farsi exhibits compensatory lengthening, which is triggered by the deletion of glottal consonants in coda position in informal speech (Darzi, 1991). As a result, minimal pairs such as [tar] and [tarh] should contrast only with respect to vowel length. A corpus of 90 words of the form CVC, CVCG, CVGC, and CVCC (where V=a vowel and G=a glottal consonant) was recorded, and durations of vowels in different contexts were measured and compared. Preliminary results show that lax vowel durations fall into three groups with CVCC longer than CVCG/CVGC, and the latter longer than CVC. It remains to be seen whether CVCG/CVGC words show compensatory lengthening when the glottal consonant is deleted.

5pSC35. Gestural stability in vowels. Thomas Purnell (Dept. of Linguist., Univ. of Wisconsin, Madison, 1168 Van Hise, 1220 Linden Dr., Madison, WI 53706, tcpurnell@wisc.edu)

In accordance with proper perception of linguistic sound units, past research has demonstrated some degree of acoustic and physiological stability. In contrast, articulatory stability has been thought to be inconsistent because articulations may vary so long as the vocal tract area function results in appropriate formant structure [Atal *et al.*, J. Acoust. Soc. Am. **63**, 1535–1555 (1978)]. However, if the area function for the constriction and its anterior region can maintain acoustic stability, articulatory stability should be observed in the relational behavior of four tongue pellets used in xray microbeam data. Previous work examined normalized pellet data in order to arrive at an average posture for each vowel [Hashi *et al.*, J. Acoust. Soc. Am. **104**, 2426–2437 (1998)]. But by assuming static (average) gestures, the research fell short of a correct postural characterization. This study of tongue pellet speed and normalized pellet displacement of front vowels spoken by ten microbeam database subjects reports that the tongue tip pellet speed maxima identify vowel edges (end of vowel onset, beginning of offset) while displacement of the three anterior pellets identify changes in formant structure (e.g., two stable regions in the Northern Cities English front low vowel).

5pSC36. Aeroacoustics production of fricative speech sounds. Michael Krane (Ctr. for Adv. Information Processing, Rutgers Univ., Piscataway, NJ 08854) and Settles Gary (Penn State Univ., University Park, PA 16802)

The aeroacoustic production of fricative speech sounds is studied using schlieren imaging and acoustic measurements. The focus is the structure of the turbulent jets formed during fricative speech sound production, and how interaction of the jet with articulators affects the acoustic nature of the speech sound. Patterns of the jets formed during the articulation of both voiced and unvoiced fricatives (s and z) are shown using schlieren images. In particular, the interaction of the jet with articulators such as teeth and lips are clearly seen, and demonstrated by varying articulator positions. Pressure measurements were made in conjunction with the images using a microphone placed near the teeth and one in the farfield. The pressure measurements show the acoustic consequences of the various jet/articulator interactions, further clarifying which articulators are most important in determining the aeroacoustic source characteristics.

5pSC37. A new taxonomy of American English /r/ using MRI and ultrasound. Mark K. Tiede (Haskins Labs, 270 Crown St., New Haven, CT 06511 and M.I.T.-R.L.E., Cambridge, MA, tiede@haskins.yale.edu), Suzanne E. Boyce, Christy K. Holland, and K. Ann Choe (Univ. of Cincinnati, Cincinnati, OH)

In this work we present preliminary results from a large scale production study of American English liquids. MRI and ultrasound have been used to image 20 subjects producing /r/ and /l/. Subjects were native speakers, representative of the main American dialects, and balanced between sexes. MRI data were collected volumetrically for recovery of

three-dimensional tongue shape. In addition, a short (1 s) midsagittal protocol was used in conjunction with ultrasound scanning to confirm the validity of the sustained production required for volumetric imaging. Our results show tongue shapes not inconsistent with the six main types found by Delattre and Freeman (1968), but their range and variety suggest that either additional canonical shapes or a different organizational principal (such as palatal shape) is motivated. We have also observed consistent patterns of tongue grooving associated with different midsagittal shapes, from the parasagittal data not accessible from the Delattre and Freeman cineradiography. In addition to this 3D-based taxonomy of /r/ tongue shapes we present similar data for lateral productions, and the associated acoustic structures for both. [Research supported by NIH.]

5pSC38. Some effects of prosodic structure on the production of /u/ in French. Marija Tabain (MARCS, UWS, Sydney, NSW 1797, Australia), Pascal Perrier, and Christophe Savariaux (ICP, INPG, Grenoble, France)

In this paper we present formant data and EMA (Carstens) data from two female speakers of French who produced /u/ in domain-final position at four different prosodic boundaries (in hierarchical order: utterance, intonational phrase, accentual phrase, and word boundaries). The prosodic boundaries are used in order to control the acoustic duration of the /u/, with greatest duration at the strongest boundary (utterance) and shortest duration at the weakest boundary (word). We present results on lip protrusion and tongue body targets with respect to these prosodically induced differences in duration, and compare these articulatory results with measurements of $F1$ and $F2-F0$ in Bark. We examine trade-offs between lip and tongue targets as duration is reduced, and speculate as to whether the speakers may have been aiming for an articulatory target or an acoustic target for /u/ at the stronger prosodic boundaries. This work combines our previous work on articulatory prosody of /a/ [Tabain, J. Acoust. Soc. Am. **113**, 516–531 (2003)] with our previous work on perturbation of /u/ [Savariaux *et al.*, J. Acoust. Soc. Am. **98**, 2428–2442 (1995)]. [Work supported by an Australian Research Council fellowship to the first author.]

5pSC39. Production and perception of Persian geminate stops at three speaking rates. Benjamin B. Hansen (Dept. of Linguist., Univ. of Texas, Calhoun Bldg. 501, Austin, TX 78712)

An experiment was designed to determine whether the geminate/singleton category distinction is maintained at fast speaking rates in Persian. Three speakers of Tehrani Persian read test words containing [t,t:,d,d:] in carrier sentences at three speaking rates. The categories do not overlap within a given speaking rate, but the fastest geminates do overlap the normal-rate singletons, implying that the listener must take speaking rate into account in order to perceive the category distinction. The ratio of the consonant closure to the preceding vowel (C/V) is not a useful rate-independent parameter for describing the geminate/singleton boundary in Persian since in Persian the vowel preceding a geminate is slightly longer. However, it was found that the marginal consonant closure (above a minimum closure of about 20 ms) maintains a fixed proportion of the average syllable duration, regardless of rate. This fixed proportion is distinct for geminates and singletons, and so may be used as a single rate-independent parameter for defining the category distinction. Perception tests on natural sentences showed that the distinction is perceptible at each of the three speaking rates. The perceptual response to manipulation of the closure durations indicated that, besides duration, additional cues to the distinction are present.